UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN

Strategy selection and strategy efficiency in mental arithmetic

Ineke Imbo

Promotor: Prof. Dr. André Vandierendonck

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologische Wetenschappen

2007

UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN

Strategy selection and strategy efficiency in mental arithmetic

Ineke Imbo

Promotor: Prof. Dr. André Vandierendonck

Proefschrift ingediend tot het behalen van de academische graad van Doctor in de Psychologische Wetenschappen

2007

CONTENTS

ACKNOWLEDGEMENTS

Dit proefschrift was mogelijk dankzij een doctoraatsmandaat van het Bijzonder Onderzoeksfonds (BOF) van de Universiteit Gent. Ik dank deze instelling voor haar steun. Daarnaast zijn er natuurlijk heel wat personen op wie ik de voorbije jaren kon rekenen.

Eerst en vooral wil ik Prof. dr. André Vandierendonck, de promotor van dit proefschrift, bedanken. Ondanks zijn erg drukke agenda vond hij altijd tijd om mijn experimentplannen te overlopen en mijn teksten kritisch te analyseren. Ook op inhoudelijk niveau stuurde hij me aan op meer diepgang. Verder dank ik ook de leden van mijn begeleidingscommissie, Prof. dr. Wim Fias, Prof. dr. Bert Reynvoet, en Prof. dr. Yves Rosseel. Bedankt voor jullie diepgaande, kritische, en positieve opmerkingen die een ware inspiratie vormden voor dit proefschrift. Yves, het was tof jou als medeauteur van een van de hoofdstukken van dit proefschrift te hebben, en bedankt dat ik telkens bij jou terecht kon voor allerhande statistische vragen. Stijn De Rammelaere, het is dankzij jou dat ik de smaak van de experimentele psychologie pas goed te pakken heb gekregen. Bedankt!

Many thanks go to Prof. dr. Patrick Lemaire and to dr. Sandrine Duverne, the persons who guided me through a nice and instructive training period in the sunny South of France. There, at the Université de Provence in Marseille, I not only learned how to implement the choice/no-choice method; I also enjoyed the French climate, food, and cordiality. Thanks!

Onderzoek doe je natuurlijk niet alleen. Ik ben blij dat ik hier op de Vakgroep Experimentele Psychologie kon rekenen op heel wat collega's. Bedankt om me te helpen bij inhoudelijke vragen (Jolien en Maud), bij het programmeren (Filip en WimN), bij wayfinding problems in Italië (Baptist), bij corsi- en synesthesieonderzoek (Arnaud en WimG), bij de studieondersteuning (Frederick), en bij de lay-out van dit proefschrift (Wouter).

Verder dank ik Antoine, Lies, en Liliane voor de technische en praktische ondersteuning; en in het bijzonder Pascal en Christophe, die steeds paraat stonden om voicekey- en computerproblemen op te lossen en zodoende mijn stressniveau onder de drempelwaarde te houden.

Tijdens de voorbije jaren had ik het plezier mijn bureau te kunnen delen met twee studiegenoten, die er allebei enorm goed in geslaagd zijn mij blij te verrassen. Eerst zat ik samen met Isabel, die na een jaartje de onderzoekswereld vaarwel zei, haar droom achterna ging, en een nieuw leven begon in café De Gekko. Gelukkig zat ik niet lang alleen, want een paar verhuistrajecten later kwam Jolien mij vergezellen. Jolien, ik heb veel van je geleerd en vond het echt leuk samen met jou een bureau te delen. Je hebt me wel goed liggen gehad met je maagontsteking, die er gelukkig geen bleek te zijn ;-) Ik ben ervan overtuigd dat je niet alleen een goede onderzoekster bent, maar ook een fantastische mama voor je dochtertje Lena.

Inspirerende collega’s vond ik niet enkel in Gent, maar ook aan de universiteit van Leuven. Bert De Smedt, Joke Torbeyns en Koen Luwel, dankjewel voor al jullie hulp tijdens de voorbije jaren. Als experts in ontwikkelingsonderzoek en de choice/no-choice methode heb ik vaak op jullie kunnen rekenen. Hoewel onze communicatie veeleer van virtuele aard was, was het telkens leuk jullie ook *in real life* te ontmoeten op talrijke congressen in België en daarbuiten.

Maar (gelukkig maar?) het leven is meer dan doctoreren alleen. Ik heb het geluk vrienden te hebben die ware experts zijn in het doen-vergeten-van-datgene-waar-ik-ook-alweer-mee-bezig-was. Dankjewel, Sarah voor de heerlijke, supergezellige etentjes en feestjes chez Marino, voor de ontspannende en inspirerende gesprekken; kortom, voor alle steun. Dankjewel, An voor de ontspannende yoga belevenissen en de stoere *char à voile* verhalen. Dankjewel Michael en Nadine, om me de Duitse gastvrijheid, keuken, en taal te doen ontdekken. Het bezoek aan de chocoladefabriek was echt een droom die werkelijkheid werd. Joke en Gwen, ook al zagen we elkaar de laatste tijd niet zo vaak, ik weet dat ik altijd op jullie kan rekenen. Bedankt!

Zoals reeds gezegd, was doctoreren niet mijn enige passie. Om af en toe de batterijen op te laden was er niets beters dan de natuur in te trekken, de bergen op en af te crossen, en nieuwe culturen op te snuiven. Ik heb het grote geluk vrienden te hebben die ook aan dergelijke reiskriebels lijden. Dankjewel, Bart en Elke, Bart en Els, Griet en Lieven, Katrien en Wouter, Karen, Steven, Riet, en Sam. Jullie zijn echt fantastische vrienden én reisgenoten; de trektochten doorheen de Spaanse Pyreneeën, de Italiaanse Alpen, en het Marokkaanse Atlasgebergte zouden hetzelfde niet geweest zijn zonder jullie. Els, Griet, Katrien, Steven en Laura, ik vind het fantastisch dat jullie me keer op keer overtuigd hebben om toertjes te gaan lopen. Vooral *achteraf* deed het telkens zo'n deugd…

Een heel grote dankjewel gaat uiteraard ook naar mijn ouders, die me steeds onvoorwaardelijk gesteund hebben in alles wat ik deed. Mama, bedankt om van ons maandelijks bezoek in Koksijde telkens een feest te maken. Papa, bedankt voor de o-zo deugddoende GR wandeltochten. Een grote dankjewel ook voor mijn broer Martijn, en voor mijn lieve meter, die me al te vaak verwende.

Tot slot wil ik een persoon bedanken die ik nog niet kende toen ik begon aan dit proefschrift. Ik was amper een paar maand bezig toen ik hem ontmoette en hij mijn leven overhoop gooide: Steven, mijn vriend, mijn lief, mijn soulmate, mijn partner-in-crime. Dankjewel om me te inspireren met je gekke gedachten en je originele ideeën. Dankjewel om te steunen, vooral die laatste maanden, waar je me geholpen hebt overal waar je kon. Hoe kan ik je beter bedanken dan met de volgende woorden: Dankjewel, Sorry, en Ik zie je graag.

Ineke, januari 2007

CHAPTER 1
STRATEGY USE IN MENTAL ARITHMETIC: AN OVERVIEW

"Can you do addition?" the White Queen asked.

"What's one and one and one and one and one and one and one and one and one and one?"

"I don't know," said Alice. "I lost count."

Lewis Carroll

Mental arithmetic is not only a key component of children's elementary education, it is also important in each adult's daily life. The knowledge of elementary arithmetic (i.e., being able to solve simple addition, multiplication, subtraction, and division problems) is a pervasive requirement of everyday modern life, providing essential means for dealing with a diverse variety of problem-solving situations. Basic arithmetic also provides the foundation for more advanced mathematical skills that are central to all modern scientific disciplines. Consequently, understanding this fundamental intellectual skill is an important goal for cognitive science. This can also be said, though, about other cognitive skills such as language and memory. However, there is something special about arithmetic that makes it different from all other skills…

The mental-arithmetic skill is unique in that you do not need to store all arithmetic facts in your long-term memory. Indeed, all simple-arithmetic problems can be calculated as well as being retrieved from long-term memory. In that respect, arithmetic problem solving importantly differs from problem solving in other cognitive domains. For example, you cannot work out the name of the capital of Papua-New-Guinea if you don't know it; neither can you translate the Dutch word 'hoofdrekenen' into the English words 'mental arithmetic' if you have not learned and stored these words before. However, you can always work out the problem 7 + 4. If you did not store the correct solution in your long-term memory, you may count until you reach the solution: 7 … 8 … 9 … 10 … 11.

This unique characteristic of mental arithmetic has only become a topic of investigation in the more recent years. Indeed, most studies in the past few decades investigated arithmetic cognition *in general*, with as main topic the structure and organization of people's mathematical knowledge in long-term memory. The question as to how this knowledge is accessed and applied in various settings has only been studied in the more recent years. This doctoral dissertation deals with the online processes in people's mind when they are solving simple-arithmetic problems, and more specifically with people's arithmetic *strategy* use. People's arithmetic strategy use entails two components: strategy selection (occurring before a particular strategy is executed) and strategy execution (occurring when a particular strategy is used to solve the arithmetic problem). The execution of strategies is often examined in terms of strategy efficiency, which refers to the speed and accuracy with which strategies are implemented.

One of the more interesting findings from this research on arithmetic strategies is that the use of nonretrieval strategies (also called procedural strategies) is not restricted to children's arithmetic problem solving. Indeed, adults had for long time been assumed to use nothing but retrieval strategies. However, normally developed and highly educated adults still use nonretrieval strategies to solve simple-arithmetic problems (e.g., LeFevre, Sadesky et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). This finding has made it necessary to re-examine some of the empirical effects in basic mathematics research in order to clarify and qualify them, so that more accurate models of mathematical cognition could be developed.

This introduction section (Chapter 1) provides the reader first with a glance on the history of mathematical models and their relevance for arithmetic strategy studies. Next, we discuss some general pieces of information about arithmetic strategy use. A short overview of some methodological issues relevant to the studies reported in this thesis is provided as well. Then, a non-exhaustive overview of the literature on four issues relevant in mental arithmetic, and highly relevant to people's arithmetic strategy use, is presented. First, what are the effects of experience and practice on adult's simple-arithmetic performance? Second, what is the role of working memory in simple-arithmetic performance? Third, how does simple-arithmetic performance develop? And finally, which individual differences affect simple-arithmetic performance? Whereas most of these questions can readily be answered in *general* terms, it is rather unclear whether they can be answered with respect to the different *strategies* people use to solve arithmetic problems. Hence, the research goal of the present doctoral dissertation was to find some answers to these four questions with respect to people's arithmetic strategy use (i.e., strategy selection and strategy efficiency).

MATHEMATICAL MODELS AND THEORIES

There are many theories about how numerical cognition is represented in memory. Each model has its own explanation for the well-known problem-size effect (also called problem-difficulty effect). This effect refers to the fact that solving large problems (e.g., 8 x 7) is slower and more error-prone than solving small problems (e.g., 3 x 4). Even more interesting for the current thesis, is that each model also makes (implicit or explicit) assumptions about which strategies are used to process numerical information. In the following, we review some relevant models of numerical cognition, with specific attention for what they say about strategy use.

The first cognitive investigation of mental arithmetic was the Groen and Parkman paper (1972), in which a counting model, the ‘min model’ was presented. According to this model, children solve simple addition problems by incrementing the largest value by ones for a total of *n* times (e.g., 3 + 5 = 5… 6… 7… 8). Although Groen and Parkman acknowledged that adults used a combination of retrieval and counting processes, their counting-based model (which supposed that the problem-size effect would increase linearly with the size of the smallest addend) appeared inadequate for explaining adults’ arithmetic performance. This urged researchers to develop retrieval-based models.

Such network-retrieval models (e.g., Ashcraft, 1982; Ashcraft & Battaglia, 1978; Geary, Widaman, & Little, 1986; Miller, Perlmutter, & Keating, 1984; Stazyk, Ashcraft, & Hamann, 1982; Widaman, Geary, Cormier, & Little, 1989; Widaman & Little, 1992; Widaman, Little, Geary, & Cormier, 1992) supposed that all arithmetic facts are stored in an interconnected network in long-term memory. When an arithmetic problem is presented, activation spreads among the number nodes in the network. The most highly activated number node is then selected as the answer to the problem. The distance - or the area - that must be searched by the spreading activation process determines the difficulty of retrieval (cf. the problem-size effect). Addition and multiplication are supposed to rely on similar cognitive processes, and their performance is supposed to be best predicted by structural variables such as the sum, the square of the sum, or the product. These models accounted reasonably well for various kinds of observed data. A disadvantage, however, was that the notion of structural variables such as the sum and the product had very limited psychological plausibility.

Another retrieval model heavily relying on structural factors is the network-interference model of Campbell (1987a, 1995; Campbell & Clark, 1989, 1992; Campbell & Graham, 1985; Campbell & Oliphant, 1992). In this model, retrieval from long-term memory is seen as a process affected by interference. A problem not only activates the correct answer but also many other, neighboring answers. Each problem thus becomes associated with a set of candidate answers, among which the retrieval process must discriminate in order to select the most strongly activated answer. The larger the interference created by competing problem-answer associations, the lower the activation level of the correct answer node, and the lower the probability of retrieval. Interference of false answers is thus an unavoidable part of the retrieval process, and especially for large problems (cf. the problem-size effect). Indeed, large problems are more similar in magnitude to their neighbors than are small problems, which results in higher interference levels for large than for small problems. A very recent version of a comparable retrieval model is the interacting neighbors model (Verguts & Fias, 2005), in which successful retrieval is related to the consistency of the answer's decade-unit digits with the decade-unit digits of close operands. Because solutions of small problems are more consistent with their neighboring answers than are solutions of large problems, the problem-size effect arises. Note that network-retrieval models and network-interference models are silent on the issue of strategies other than retrieval, and their influence on arithmetic processing.

Gradually, non-structural variables came into play. In such models (e.g., Ashcraft, 1987; Hamann & Ashcraft, 1985; see also Campbell & Graham, 1985), the strength with which number nodes are stored and interconnected was seen as a function of experiential factors such as occurrence frequency, acquisition order, and practice frequency. Because small problems are encountered and practiced more frequently than are large problems, they have stronger problem-answer associations (cf. the problem-size effect). Moreover, the early acquisition of small problems may make the acquisition of large problems more difficult (cf. proactive interference). Structural variables were said to be only coincidently important because they correlated with non-structural variables. Roughly coincidental with the discovery of the importance of non-structural variables, the research field also re-acknowledged the use of nonretrieval strategies. In the model of Ashcraft (1987), for example, solving arithmetic problems relies on a ‘horse race’ between declarative (retrieval) routes and procedural (nonretrieval) routes. The retrieval process thus occurs *in parallel* with a procedure-based solution attempt. Among children, retrieval processes are slow and they will fail, whereas in adults, the race is generally won by the yet faster retrieval processes.

Another model that explicitly includes the possibility of nonretrieval strategies is the distribution of associations model of Siegler and Shrager (1984). In this model, the basic representation of a problem is accompanied not only by the problem’s correct answer, but also by incorrect answers that the individual has generated or computed across experience. Problems have *peaked* distributions when the association between the problem and its correct answer is strong while the associations with other answers are weak. Problems have *flat* distributions of associations when the association with the correct answer is only slightly stronger than the associations with other, incorrect answers. The more peaked the distribution of associations, the higher the probability of retrieval. Due to past experiences, small problems have more peaked distributions of associations than do large problems, resulting in the well-known problem-size effect. In the distribution of associations model, nonretrieval strategies occur *after* a failed retrieval attempt. Retrieval fails when the association strength does not cross the confidence criterion (which determines how sure one must be to state a retrieved answer) or when the search length criterion (which determines how many attempts one will make to retrieve an answer before trying a nonretrieval strategy) is exceeded. The more memory searches are needed, the slower the retrieval speed. Each time an answer is stated, the association between that answer and the problem becomes stronger – regardless of whether the answer was produced through a retrieval or nonretrieval strategy.

In a later model, the adaptive strategy choice model (ASCM[1], Siegler & Jenkins, 1989; Siegler & Shipley, 1995), retrieval is no longer the default strategy for a first attempt. According to the ASCM, people accumulate information about each strategy's efficiency (i.e., speed and accuracy). This strategy information then determines the problem-strategy association strength. The model thus first selects a strategy based on the distribution of strategy strengths, and then attempts to execute that strategy. Whether one first attempts to solve the problem with a retrieval or nonretrieval strategy depends on the relative association strengths. Problems with *flat* distributions of associations generally have weak problem-answer associations and strong problem-procedure associations. Accordingly, the probability of retrieving an answer on the first retrieval attempt is small; if the answer is retrieved at all, multiple retrieval attempts will be required. If no answer exceeds the confidence criterion, a nonretrieval strategy will be used to solve the problem. Problems with *peaked* distributions of associations, in contrast, will readily be solved with the retrieval strategy. With experience, the strength of the retrieval strategy will overpower other association strengths. Consequently, retrieval will dominate arithmetic problem solving. An important difference with the so-called 'horse-race' models, where several strategies are activated simultaneously, is that in the ASCM only *one* strategy can be operative at any given moment.

In 1992, Ashcraft published an integrated model, one that incorporated characteristics of various previous models (e.g., Ashcraft, 1987; Siegler & Jenkins, 1989, Campbell, 1987a). In this associative network-retrieval model, arithmetic facts are stored in an interrelated network in long-

[1] There exist several recent adaptations of the ASCM, such as the strategy choice and discovery simulation (SCADS) model (Shrager & Siegler, 1998), which incorporates the ASCM but also models metacognitive processes to allow for the discovery of new strategies; and the SCADS* model (Siegler & Araya, 2005), which adds six new mechanisms to the SCADS model (i.e., controlled attention, interruption of procedures, verbalization, priming, forgetting, and dynamic feature detection).

term memory. Presentation of an arithmetic problem results in activation of number nodes specified in the problem. Activation then spreads along associative links, so that related nodes, such as the sum and the product, get also activated. Importantly, each problem has associations to both correct and incorrect answers. Another essential aspect of this theory is that retrieval and nonretrieval strategies are triggered *in parallel*, with the faster route governing performance. Because several strategies can be active at any given moment, trials being processed via a nonretrieval strategy may be disrupted by retrieval-based interference.

In contrast to most retrieval models, Baroody's (1983, 1984, 1994) schema-based view maintains that *nonretrieval* strategies are extremely important in both children's and adults' arithmetic performance. According to Baroody, the key change in arithmetic skill involves a shift from slow, effortful nonretrieval strategies to fast, automatic nonretrieval strategies. He denies that nonretrieval strategies are inherently slow and argues that the problem-size effect should be interpreted in terms of how well procedural knowledge is internalized and automatized. Importantly, Baroody does not deny the strengthening of problem-answer associations. Though, he asserts that semantic knowledge on number relationships is an inseparable component of people's long-term memory network. He further notes that elaborating procedural knowledge is cognitively more economic than storing all individual facts in long-term memory. Development (and mastery) of arithmetic skill should thus be seen as learning a system of rules, procedures, and principles rather than memorizing hundreds of specific numerical associations.

A final theory worth mentioning here is the triple-code theory of Dehaene (1992, 1997; Cohen & Dehaene, 2000, Cohen, Dehaene, Chochon, Lehericy, & Naccache, 2000; Dehaene & Cohen, 1995, 1997). According to this theory, there are three kinds of number codes in the human brain: visual-Arabic, auditory-verbal, and analog magnitude codes. We mention this theory because it makes specific predictions about the strategies needed to solve simple-arithmetic problems. Addition and multiplication are regarded as rote verbal memory tasks that utilize the auditory-verbal number code. Consequently, addition and multiplication problems are hypothesized to be retrieved as automatic verbal associations. Subtraction and division problems, in contrast, would be solved using the analog magnitude code, i.e., through mental manipulation of the quantities being represented, also called semantic elaboration.

ARITHMETIC STRATEGIES

Approximately 10 years ago, LeFevre and colleagues (1996a, 1996b) published data that put all retrieval-based theories and models into question. They showed that adults, just like children, use both retrieval and nonretrieval strategies to solve simple-arithmetic problems. This idea was not new, however, as earlier studies also had reported substantial amounts of nonretrieval strategies in adults' arithmetic problem solving (e.g., Geary, Frensch, & Wiley, 1993b; Geary & Wiley, 1991; Svenson, 1985).

According to LeFevre et al. (1996a, 1996b), this multiple strategy use provided new insights in the well-known problem-size effect. Indeed, as the (generally slower) nonretrieval strategies are used more frequently on large problems than on small problems, LeFevre and colleagues maintained that the problem-size effect might – to a certain extent – be caused by strategy selection processes. Accordingly, a person's simple-arithmetic performance was suggested to be best predicted by that person's percentage retrieval use. The percentage retrieval use indeed appeared to be a good predictor of retrieval latencies; moreover, this predictor was said to have a considerable psychological plausibility. Structural variables (such as the sum and the product) were said to be reliable predictors only because of their correlation with strategy selection.

The overall pattern of results observed by LeFevre and colleagues (1996a, 1996b) – including multiple strategy use, more frequent retrieval use on small problems than on large problems, higher efficiency on retrieval trials than on procedural trials, and a smaller problem-size effect on retrieval trials than on procedural trials – has since been observed in numerous adult studies (e.g. Campbell & Austin, 2002; Campbell & Fugelsang, 2001; Campbell & Gunter, 2002; Campbell, Parker, & Doetzel, 2004; Campbell & Timm, 2000; Campbell & Xue, 2001; Hecht, 1999; LeFevre & Morris, 1999; Robinson, Arbuthnott, & Gibbons, 2002; Seyler, Kirk, & Ashcraft, 2003; Smith-Chant & LeFevre, 2003).

Although the selection and the efficiency of strategies differ across operations (Campbell & Xue, 2001), the problem-size effect is significant in all four operations. It is now generally acknowledged that there are at least three factors contributing to the problem-size effect: (1) more efficient retrieval use on small problems than on large problems, (2) more efficient procedural use on small problems than on large problems, and (3) more frequent retrieval use on small problems than on large problems. More specifically, Campbell and Xue (2001) argued that about half of the problem-size effect in simple arithmetic is due to slower retrieval for large problems, and half is due to greater use of nonretrieval strategies for large problems.

The discovery of multiple strategy use further revealed lots of research questions. On the *theoretical* level, it was argued that models of mathematical cognition should incorporate a link between nonretrieval strategy use and arithmetic performance (e.g., LeFevre et al. 1996a, 1996b; see also Baroody, 1994). As noted before, early models of arithmetic problem solving were mainly interested in the long-term memory structure and organization of simple-arithmetic facts (Ashcraft, 1992; Campbell & Oliphant, 1992; Widaman et al., 1989; Widaman & Little, 1992). As these models depart from the storage of simple-arithmetic facts in an interrelated long-term memory network, they often (implicitly) predict that adults rely largely if not exclusively on direct retrieval. In the Discussion section of this thesis (cf. Chapter 8), multiple-strategy models are reviewed and discussed.

On the *empirical* level, the discovery of multiple strategy meant that all well-known effects described in the literature on mental arithmetic should now be re-investigated with regards to *strategic* processes. Indeed, many differences in the performance between two individuals and in the performance of a single individual under different situations may be accounted for by differences in their underlying cognitive processes, i.e., their strategies. Consequently, we need to understand differences in strategy use *within* as well as *between* individuals (Lemaire & Fabre, 2005). In the current doctoral dissertation, three rather well-known effects observed in arithmetic studies are re-investigated with regard to people's strategy selection and strategy efficiency: the role of practice, the role of working memory, and developmental trajectories. The impact of individual differences was investigated as well. Before considering each of these effects, some methodological issues relevant to the studies presented in this thesis are discussed.

SOME METHODOLOGICAL ISSUES[2]

In the current doctoral dissertation, we decided to study simple-arithmetic problems only. Among addition and multiplication, simple-arithmetic problems refer to one-digit problems, i.e., ranging from 1 + 1 to 9 + 9 for addition and ranging from 1 x 1 to 9 x 9 for multiplication. Among subtraction and division, simple-arithmetic problems refer to the counterparts of the problems used in addition and multiplication,

---

[2] In this section, we only discuss methodological issues that are relevant to *all* experiments reported in this doctoral dissertation. One methodological issue – the *choice/no-choice method* (Siegler & Lemaire, 1997) – is relevant to only three of our studies (cf. Chapters 4, 5, and 7) and will be explained there.

respectively, i.e., ranging from 2 – 1 to 18 – 9 for subtraction and ranging from 1 : 1 to 81 : 9 for division. Complex problems involve multiple digits (e.g., 53 + 74, 12 x 69) and generally involve a whole series of cognitive processes (Geary, 1994; Geary & Widaman, 1987), such as breaking down the problem, fact retrieval, short-term storage of partial results, mental manipulation of these partial results, et cetera. This makes the empirical study of strategies more complicated (but see Chapter 8, for a discussion on complex-arithmetic strategies). Within the range of simple-arithmetic problems, we decided to exclude three types of problems, more specifically, zero-problems (e.g., 0 + 5), one-problems (e.g., 1 x 8), and tie-problems (e.g., 3 + 3; 6 x 6). Zero- and one-problems are often solved with rules (e.g., 0 + N = N and 1 x N = N), which makes the study of strategies rather difficult. Hence, if such problems are tested, they should be analyzed separately from all other problems (e.g., Campbell & Xue, 2001; LeFevre et al., 1996a, 1996b). Tie-problems were excluded because they are supposed to be coded differently in long-term memory (e.g., LeFevre, Shanahan, & DeStefano, 2004; Campbell & Gunter, 2002).

We also decided to study people's arithmetic processing by means of a production task (e.g., 7 x 8 = ?) rather than by means of a verification task (e.g., 7 x 8 = 64, true/false?; but see Chapter 3, in which both tasks were used). This decision was based on the fact that the verification of arithmetic problems poses several problems. First, the verification task is generally viewed as a four-stage process of encoding, retrieval and/or calculation, comparison to the presented answer, and response execution. In contrast, production is generally thought to entail only encoding, retrieval and/or calculation, and response execution (e.g., Romero, Rickard, & Bourne, 2006). Note that the response execution process in the verification task entails extra decision processes regarding which button to press. Second, the absence of differences in raw verification latencies does not mean that computation processes are unaffected. Indeed, people may speed up decision processes as a result of additional time spent on computation processes (Kaye, de Winstanley, Chen, & Bonnefil, 1989). Third, verification strategies are different from the strategies used in production trials (e.g., Campbell & Tarling, 1996; Cornet, Seron, Deloche, & Lories, 1988; Krueger & Hallford, 1984). For example, participants might rely on estimation or plausibility judgments to determine whether the stated answer is correct, rather than actually determining the answer to the problem. If the stated answer is not approximately correct in terms of magnitude (e.g., 3 x 8 = 72, true/false?), participants might even not start to calculate (Ashcraft & Stazyk, 1981). Participants might also use odd-even rules in order to verify quickly whether an answer is correct or not (Krueger, 1986; Lemaire & Fayol, 1995), or compare the equation as a whole and use this comparison to evaluate whether the statement is true or false (Zbrodoff & Logan, 1990). Finally, presenting the correct answer may have a direct impact on the retrieval process, which makes the verification task inherently limited in its capacity to measure retrieval efficiency (Campbell, 1987b).

A final methodological issue that is relevant with respect to all our experiments is the use of verbal strategy reports. Asking people to report which strategy they used to solve a problem has successfully been used to study children's arithmetic processing (e.g., Bisanz, Morrison, & Dunn, 1995; Carr & Davis, 2001; Cooney et al., 1988; Davis & Carr, 2002; De Smedt et al., 2006; Geary, 1996; Geary & Brown, 1991; Geary, Brown, & Samaranayake, 1991; Geary, Fan, & Bow-Thomas, 1992; Geary, Hamson, & Hoard, 2000a; Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Hecht, Close, & Santisi, 2003; Jordan & Montani, 1997; Kerman & Siegler, 1997; Lemaire & Lecacheur, 2002; Lemaire, Lecacheur, & Farioli, 2000; Mabbott & Bisanz, 2003; Noël, Seron, & Trovarelli, 2004; Robinson, 2001; Robinson et al., 2006; Siegler, 1987, 1988a, 1989; Siegler & Stern, 1998; Svenson & Sjöberg, 1982, 1983; Torbeyns, Verschaffel, & Ghesquière, 2002, 2004a, 2004b, 2005a) and to study adults' arithmetic processing (e.g., Campbell & Austin, 2002; Campbell & Fugelsang, 2001; Campbell & Penner-Wilger, 2006; Campbell & Xue, 2001; Compton & Logan, 1991; Geary et al., 1993b; Geary & Wiley, 1991; Green, Lemaire, & Dufau, in press; Hecht, 1999; Hoyer, Cerella, & Onyper, 2003; Kalaman & LeFevre, in press; LeFevre et al., 1996a, 1996b; Logan & Klapp, 1991; Onyper, Hoyer, & Cerella, 2006; Robinson et al., 2002; Romero et al., 2006; Seyler et al., 2003; Smith-Chant & LeFevre, 2003; Svenson, 1985; Touron, Hoyer, & Cerella, 2006).

In our experiments, we asked participants – after each trial – to classify their strategy into one of four categories: Retrieval, Transformation, Counting, and Other. These strategy categories have been frequently used in previous studies and have been shown to account well for the observed data (e.g., Campbell & Austin, 2002; Campbell, Fuchs-Lacelle, & Phenix, 2006; Campbell & Fugelsang, 2001; Campbell & Gunter, 2002; Campbell & Penner-Wilger, 2006; Campbell & Timm, 2000; Campbell & Xue, 2001; Campbell et al., 2004; Geary & Wiley, 1991).

It should be noted, though, that the use of verbal strategy reports has been criticized (e.g., Kirk & Ashcraft, 2001; Russo, Johnson, & Stephens, 1989). Being aware of these criticisms, we tried our best to obtain unbiased and valid information about people's strategy choices. It has been shown that participants are able to report thought sequences retrospectively, provided that the task is of relatively short duration and the reports are gathered immediately after the task is completed (Ericsson & Simon, 1980, 1984, 1993). A proper use of verbal strategy reports further entails completely unbiased instructions of the experimenter. It also precludes response deadlines because fast response deadlines increase reported retrieval use (Campbell & Austin, 2002). All these suggestions were taken into account in our experiments. The discussion on the merits and limits of verbal strategy reports continues in the last chapter of this thesis (cf. Chapter 8), where we also provide some alternative methods.

In all our experiments, we combined self-report data with latency data. That is to say, trials were first separated by self-reports, and then latencies were analyzed. This labor-intensive methodology does generally not allow for large cohorts of participants to be studied (Hopkins & Lawson, 2002). However, it is one of the most successful ways to yield very accurate information about *which* strategies are applied and *how* the strategies are applied. Indeed, averaging latencies across strategies can result in very misleading conclusions about people's problem-solving processes (e.g., Cooney, Swanson, & Ladd, 1988; Siegler, 1987, 1989).

EXPERIENCE-RELATED EFFECTS

Without doubt, people differ greatly in their efficiency in solving arithmetic problems. What causes these inter-individual differences? In the first two empirical chapters (Chapters 2 and 3), we investigated to which extent arithmetic experience continues to affect people's strategy use once the elementary-school period lays already far behind.

In many mathematical models, experiential factors are the main determinants of mental representation, acquisition, and subsequent performance (e.g., Ashcraft, 1987; Campbell & Graham, 1985; Siegler, 1988b; Siegler & Shipley, 1995). However, studies comparing the simple arithmetic performance of adults with different experiences started only about 10 years ago (e.g., Geary, 1996; Geary, Salthouse, Chen, & Fan, 1996b; LeFevre & Liu, 1997). Most of these studies compared the arithmetic performance across cultures with different educational and linguistic experiences (e.g., Canadian adults vs. Chinese adults). In the current doctoral dissertation, we wondered whether – within one single culture – reasonably skilled adults might also have different experiential backgrounds, and, accordingly, different arithmetic performance patterns.

We investigated practice effects on adults' simple-arithmetic performance implicitly in Chapter 2 and explicitly in Chapter 3. More specifically, in Chapter 2, the effects of daily arithmetic experience were studied. This variable was rather ecological, as it incorporated long-lasting experiences and was not experimentally manipulated. In Chapter 3, this limit was overcome by testing practice effects experimentally. In the following paragraphs, we provide an overview of empirical evidence on both types of practice. As one may notice, many studies interpreted practice effects in terms of strategy selection and/or strategy efficiency related processes. However, only very few of them really included strategy reports. In contrast to most previous practice studies, we included strategy reports, which enabled us to infer whether practice influenced strategy selection, strategy efficiency, or both.

EXPERIENCE-BASED PRACTICE

Although the evidence about experience-based practice effects on adults' arithmetic performance is very limited, three studies are worth mentioning here. The first one is Hecht's (1999) study, in which effects of math achievement on adults' simple addition and multiplication performance were investigated. The math achievement tests were supposed to depend largely on the participants' history of math course work. Percentages retrieval use, retrieval speed, and retrieval accuracy were higher for adults with higher levels of math achievement than for adults with lower levels of math achievement. Roussel, Fayol, and Barrouillet (2002) compared experienced participants' (primary school teachers) and inexperienced participants' (undergraduate psychology students) arithmetic performance. The latter group performed significantly slower than the former, especially on large problems. A final real-live investment of practice effects has recently been reported by Verschaffel, Janssens, and Janssen (2005). These researchers tested adults' mathematical competences before and after their 3-year training to become an elementary-school teacher. At the end of the 3-year training, the overall test performance had become substantially better. In the first study presented in this thesis (cf. Chapter 2), we investigated whether the experience-based measure 'daily arithmetic practice' predicted young adults' strategy efficiency and/or strategy selection processes.

EXPERIMENTALLY-MANIPULATED PRACTICE

Explicit practice effects on arithmetic performance have been investigated quite frequently. Explicitly practicing simple-arithmetic problems increases people's performance on these problems. Smaller latencies on simple-arithmetic tasks after practice than before have been observed in adults (e.g., Campbell, 1987a, 1999; Fendrich, Healy, & Bourne, 1993; Pauli, Bourne, & Birbaumer, 1998; Rickard, 2005; Rickard & Bourne, 1996; Rickard, Healy, & Bourne, 1994; Royer, Tronsky, Chan, Jackson, & Marchant, 1999a) as well as in children (e.g., Goldman, Mertz, & Pellegrino, 1989; Goldman, Pellegrino, & Mertz, 1988). Practice effects have also been shown in complex-arithmetic tasks (e.g., Spelke & Tsivkin, 2001), in pseudo-arithmetic tasks (e.g., Rickard, 1997; Onyper et al., 2006; Touron et al., 2004), and in the alphabet arithmetic task (B + 3 = E, e.g., Brigman & Cherry, 2002; Compton & Logan, 1991; Hoyer et al., 2003; Klapp, Boches, Trabert, & Logan, 1991; Logan, 1988; Logan & Klapp, 1991; Zbrodoff, 1995, 1999). Such practice-related increases in speed have generally been hypothesized to be caused by the replacement of procedural strategies by the retrieval strategy, although the evidence based on strategy reports was lacking[3]. Indeed, faster responses after practice than before may not only be caused by more frequent retrieval use, but also by more efficient strategy execution. Using strategy reports in a practice study (cf. Chapter 3) enabled us to differentiate between practice effects on strategy selection and practice effects on strategy efficiency.

---

[3] It should be noted that some studies (e.g., Compton & Logan, 1991; Rickard, 1997) included strategy reports on *subsets* of trials, e.g., on one sixth of the trials. Logan and Klapp (1991) asked participants – at the end of the experiment – to estimate the percentage of trials on which they had used retrieval vs. counting strategies. Strategy reports on *all* trials have been used in alphabet arithmetic tasks (e.g., Hoyer et al., 2003) and in pseudo arithmetic tasks (e.g., Onyper et al., 2006; Touron et al., 2004).

An important observation in most practice experiments is that the problem-size effect decreases with practice, both in ‘pure’ arithmetic tasks (e.g., Fendrich et al., 1993; Pauli et al., 1998; Rickard & Bourne, 1996) and in the alphabet arithmetic task (e.g., Brigman & Cherry, 2002; Logan, 1988; Zbrodoff, 1995). However, practice never eliminated the problem-size effect, even at asymptotic response times. The strategy-related approach, applied in the current doctoral dissertation, enabled us to test size-related effects for retrieval and nonretrieval trials separately. The disappearance of the problem-size effect would indicate that the persistent problem-size effect observed in previous studies was caused by the accidental use of nonretrieval strategies, even after extensive practice. The persistence of the problem-size effect would suggest that there is something inherently different about large problems that causes response latencies to be slower.

A final important observation is that practice effects transfer only to very similar problems (e.g., operand-reversed problems) but not to new problems (e.g., Fendrich et al., 1993; Pauli et al., 1994, 1998; Rickard & Bourne, 1996; Rickard et al., 1994). Whether or not practice transfers to related problems across operations (e.g., from 4 x 8 to 32 : 4) is still a matter of debate (see e.g., Campbell, 1999; Rickard & Bourne, 1996; Rickard et al., 1994). The lack of transfer effects has been seen as evidence for the fact that skilled arithmetic performance reflects direct memory retrieval rather than nonretrieval strategy use (e.g., Logan & Klapp, 1991). Otherwise stated, previous studies hypothesized that the beneficial effects of practice were problem specific (i.e. only for retrieval), with no role for nonretrieval strategy use. As strategy reports were included in the current study (Chapter 3), it was possible to differentiate whether transfer effects (if apparent) were related to retrieval and/or nonretrieval strategy use.

THE ROLE OF WORKING MEMORY

In Chapters 4 and 5, we present a total of four experiments in which the role of working memory in simple-arithmetic strategy use is studied. In the following paragraphs, we first provide a description of the working-memory model of Baddeley and Hitch (1974). We used this working-memory model as a conceptual framework rather than as a to-be-tested theory. Indeed, it was beyond the scope of the current thesis to test the validity of this or another working-memory theory. After the description of the working-memory model, we give a non-exhaustive overview of studies which investigated the role of working memory in simple-arithmetic problem solving. Finally, we discuss the limited number of studies which investigated the role of working memory in relation to arithmetic *strategy* use.

THE WORKING-MEMORY MODEL

Working-memory processes generally refer to the temporary storage and processing of information in a variety of tasks (Baddeley, 1986, 1990). Many models and theories about humans' working memory have been proposed (see Miyake & Shah, 1999, for an overview). Some define working memory as a unitary system that is primarily involved in attentional control (e.g., Anderson, Reder, & Lebiere, 1996; Cowan, 1988, 1995, 1999; Daneman & Carpenter, 1980; Engle, Cantor, & Carullo, 1992; Engle, Kane, & Tuholski, 1999a; Just & Carpenter, 1992; Kane, Bleckley, Conway, & Engle, 2001; Lovett, Reder, & Lebiere, 1999; Tuholski, Engle, & Baylis, 2001) whereas others define working memory as a multi-componential system composed of subsystems that are specialized to handle different kinds of information (Baddeley, 1986, 1996, 2000; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). As noted by DeStefano & LeFevre (2004), the vast majority of empirical work on working memory and mental arithmetic has been done within the multicomponential model proposed originally by

Baddeley and Hitch (1974). Accordingly, this model was used to frame the current doctoral dissertation as well. This working-memory model has several advantages. For example, there exists a large variety of secondary tasks that tax specific working-memory components. Consequently, very specific predictions can be made about the role of each specific working-memory component.

The multi-componential working-memory model of Baddeley and Hitch (1974; Baddeley, 2000) consists of four components: the central executive, the phonological loop, the visuo-spatial sketchpad, and the episodic buffer. The executive working-memory component is a modality-free, limited-capacity system that takes care of control processes, monitoring, response selection, planning and sequencing. The central executive also controls the coordination and integration of information of the slave systems. Several authors tried to fractionate the central executive into a limited amount of functions (e.g., Baddeley, 1996; Miyake et al., 2000). In the present thesis, though, we investigated the role of central executive working-memory resources *in general*. The discussion on the possible roles of the different executive functions is postponed to Chapter 8.

The phonological working-memory component can be divided in two subcomponents: an active subvocal rehearsal process and a passive phonologically based store (e.g., Baddeley, 1992; Baddeley & Logie, 1992; Logie & Baddeley, 1987). Phonological information is held in the phonological store. Because the contents of the phonological store are subject to decay, they have to be refreshed by the rehearsal process. This rehearsal process can be seen as some form of subvocal articulation, closely linked with the speech production system.

The visuo-spatial working-memory component functions as a mental blackboard or workspace for temporary storage of visual and spatial information (Logie, 1995). It has been suggested that this slave system can be subdivided in a passive (visually based) and an active (spatially based) subsystem. In the present doctoral dissertation, it was decided not to study the role of the visuo-spatial sketchpad in simple-arithmetic problem solving. First, the theoretical basis upon which a role for the visuo-spatial sketchpad in mental arithmetic might be expected is rather scarce. Second, the empirical results concerning visuo-spatial load effects on arithmetic performance are equivocal (DeStefano & LeFevre, 2004). Finally, the availability of pure visual or spatial secondary tasks is lacking: many visual tasks rely on spatial resources and vice versa (e.g., Pickering, 2001). Moreover, visuo-spatial tasks often require executive working-memory resources (e.g., Martein, Kemps, & Vandierendonck, 1999; Phillips & Christie, 1977; Wilson, Scott, & Power, 1987), which makes it difficult to differentiate between the need for visuo-spatial and executive working-memory resources.

Finally, the episodic buffer is a limited-capacity system that binds information from the subsidiary systems and from long-term memory into a unitary episodic representation (Baddeley, 2000). Given the small amount of empirical information on this working-memory component, it was decided not to study the role of this component either.

Two methods for testing the involvement of working memory are in general use. First, in the selective interference paradigm, a dual-task methodology is used. More specifically, performance on a primary task (e.g., mental arithmetic) is examined while participants perform a concurrent secondary task. This secondary task places demands on one specific working-memory component. When both tasks load the same working-memory component, concurrent task execution should decrease performance on one of both tasks (or on both tasks). Second, in the individual-difference approach, participants are given a working-memory span assessment and are then tested on the task of interest (e.g., mental arithmetic). Performance differences on the task of interest may then be interpreted as due to differences in working-memory capacity assessed by the span test.

THE ROLE OF WORKING MEMORY IN SIMPLE ARITHMETIC

It is not within the scope of this thesis to review all studies on the role of working memory in mental arithmetic. Recently, a thorough review has been published on this topic (DeStefano & LeFevre, 2004). We have limited our overview to the relevant topics, namely (a) the role of phonological working-memory resources and executive working-memory resources, rather than the role of visuo-spatial working-memory resources, and (b) the role of working memory in simple arithmetic rather than in complex arithmetic.

In one of the earliest experiments on the role of working memory in simple arithmetic (Kaye et al., 1989), participants had to verify simple addition problems under no-load and load conditions. In the latter conditions, participants had to respond to an auditory tone that was presented across the various processes needed in the arithmetic task. Response latencies were longer when the tones occurred at the same time as the problem or the answer, which led the authors to conclude that executive working-memory resources were required to process simple-arithmetic problems. Later studies – using neater secondary tasks loading the central executive – confirmed that executive working-memory resources are needed in the verification of simple addition and multiplication problems (e.g., Ashcraft, Donley, Halas, & Vakali, 1992; De Rammelaere, Stuyven, & Vandierendonck, 1999, 2001; Lemaire, Abdi, & Fayol, 1996). The phonological loop, in contrast, would not be needed in the verification of simple addition and multiplication problems (De Rammelaere et al., 1999, 2001; but see Lemaire et al., 1996). Concerning problems in which answers to simple addition and multiplication problems had to be *produced* rather than verified, the same conclusion was obtained, i.e., involvement of executive working-memory resources but not of phonological working-memory resources (e.g., De Rammelaere & Vandierendonck, 2001; Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Muyllaert, 2006; Seitz & Schumann-Hengsteler, 2000, 2002; but see Lee & Kang, 2002). More recently, it has been shown that solving simple subtraction and division problems relies on executive working-memory resources as well (e.g., Deschuyteneer, Vandierendonck, & Coeman, 2007).

Based on these studies, and on dual-task studies investigating the role of working memory in complex-arithmetic problem solving (e.g., Fürst & Hitch, 2000; Heathcote, 1994; Hitch, 1978; Imbo, Vandierendonck, & De Rammelaere, in press e; Imbo, Vandierendonck, & Vergauwe, in press g; Logie, Gilhooly, & Wynn, 1994; Trbovich & LeFevre, 2003; Widaman et al., 1989), authors tended to infer the specific roles of executive and phonological working-memory resources. The phonological loop has been hypothesized to retain initial problem information, interim solutions, and calculation accuracy. However, the phonological loop would not be needed in arithmetic fact retrieval. Central executive working-memory resources have been hypothesized to retrieve, select, and manipulate number information, to perform calculations, to sequence calculation steps, to update information available in memory, and to provide approximately correct solutions.

THE ROLE OF WORKING MEMORY IN SIMPLE-ARITHMETIC STRATEGIES

It is important to note that all studies mentioned above averaged response latencies (or other performance measures) across strategies. This can be heavily misleading if execution of only some strategies would require substantial working-memory resources, or if the role of some working-memory resources would differ across strategies. In the current doctoral dissertation, we re-open the question of working-memory involvement in adults' simple-arithmetic performance. The inclusion of strategy reports enabled us to segregate trials by strategy. To our knowledge, only two simple-arithmetic studies combined dual-tasking with strategy reports; these are reviewed below.

In Hecht (2002), participants had to verify simple addition problems under phonological and executive working-memory loads. An executive working-memory load disrupted processing for all strategies (retrieval, transformation, and counting), but only counting trials were affected by a phonological load. However, Hecht contributed the load effects on the retrieval strategy to overall processing costs associated with the verification task (e.g., comparison and decision processes). Concerning strategy efficiency, it was thus concluded that both executive and phonological working-memory resources were needed in counting, whereas no substantial phonological or executive working-memory resources would be needed in retrieval. The strategy selection process was not influenced by the availability of working-memory resources either.

In four experiments, Seyler et al. (2003) investigated adults' performance on simple subtraction problems. In the first two experiments, they observed that both response latencies and errors increased significantly for subtraction problems with a minuend of 11 or larger (e.g., 14 – 8). In the third experiment, strategy reports were recorded. Because the use of nonretrieval strategies also jumped at minuend 11, Seyler et al. concluded that the problem-size effect found in the first two experiments was largely due to the use of slow, effortful nonretrieval strategies. In the fourth experiment, the simple subtraction problems had to be solved under working-memory load conditions. The working-memory load effects increased for subtraction problems with minuends of 11 or greater. Although Seyler et al. never combined dual-tasking with strategy reports in one single experiment, combining the results of all experiments suggested that the use of nonretrieval strategies was strongly associated with greater working-memory involvement.

In the current thesis, we tested the role of working memory in simple-arithmetic strategies by combining strategy reports with various working-memory load conditions. This methodology enabled us to infer the role of working memory in strategy selection as well as in strategy efficiency. In

Chapters 4 and 5 we investigated the role of executive and phonological resources in adults' strategy use, and in Chapter 7 we investigated the role of executive resources in children's strategy use.

DEVELOPMENTAL-RELATED EFFECTS

As noted above, mental arithmetic is a key component of children's elementary education. According to the Flemish school curriculum, addition is taught in the 1st grade and multiplication is taught in the 2nd grade. Counting is thus one of the earliest interactions with numbers. Once children master the counting process (needed in addition), they can start to develop further arithmetic skills, such as subtraction (inversed counting) and multiplication (repeated counting). In the current doctoral dissertation (cf. Chapters 6 and 7), we investigated children's strategy use in solving simple-arithmetic problems, starting from 2nd grade and reaching till 6th grade. In the following, some previous developmental studies will be discussed. First, we review some basic information about children's mental arithmetic development[4]. Second, this information is extended towards strategy use. Note that, whereas multiple strategy use was seen as a rather 'surprising' observation in adults, it has always been acknowledged as common in children (Fuson, 1982; Geary, 1996; Lemaire & Siegler, 1995; Robinson, 2001; Siegler, 1987, 1988a, 1988b). Third, the role of working memory in children's arithmetic performance is reviewed.

MENTAL ARITHMETIC THROUGHOUT DEVELOPMENT

Older children perform arithmetic problems faster and more accurately than younger children do. This has been shown for simple addition problems

---

[4] We will only discuss children's arithmetic abilities from the moment on which they enter elementary school (i.e., 1st graders). For an overview on development of arithmetic abilities in younger children (infants), see e.g., Butterworth (2005).

(e.g., Adams & Hitch, 1997; Ashcraft & Fierman, 1982; Hamann & Ashcraft, 1985), simple multiplication problems (e.g., Butterworth, Marchesini, & Girelli, 1999; Campbell & Graham, 1985; De Brauwer, Verguts, & Fias, 2006; Kaye et al., 1989; Koshmider & Ashcraft, 1991), and mathematical word problems (e.g., Swanson, 2004; Swanson & Beebe-Frankenberger, 2004). Accordingly, the problem-size effect also decreases with age. Based on these results, most developmental studies supposed a transition from counting-based performance (i.e., procedural knowledge) in young children to retrieval-based performance (i.e., declarative knowledge) in older children and adults (e.g., Ashcraft, 1982; Ashcraft and Fierman, 1982; Groen & Parkman, 1972; Hamann & Ashcraft, 1985; Koshmider & Ashcraft, 1991). More specifically, children's associative network of problem-answer associations would evolve through development (Campbell & Graham, 1985) and would be established already during the early phases of skill acquisition (Cooney et al., 1988). Consequently, long-term memory knowledge would become more and more automatically activated, as confirmed by more frequent interference effects with growing age (e.g., Hamann & Ashcraft, 1985; Koshmider & Ashcraft, 1991; LeFevre, Bisanz, & Mrkonjic, 1988; LeFevre, Kulak, & Bisanz, 1991; Lemaire, Barrett, Fayol, & Abdi, 1994). However, none of the studies mentioned above included trial-by-trial strategy reports; evidence for the transition from nonretrieval strategies to retrieval was thus never made explicit.

ARITHMETIC STRATEGIES THROUGHOUT DEVELOPMENT

The transition hypothesis postulated above (i.e., the transition in children's strategy use from counting to retrieval) has been confirmed by developmental studies including strategy reports. The earliest studies on children's strategy choices are reported by Svenson and colleagues (Svenson, Hedenborg, & Lingman, 1976; Svenson & Sjöberg, 1981; 1982, 1983). In these studies, children had to solve simple addition and subtraction problems, and to provide trial-by-trial strategy reports. Across age, retrieval use increased whereas counting use decreased, confirming the transition hypothesis. Comparable results were obtained by Carpenter and Moser (1984) with simple addition and subtraction word problems. However, because none of these studies recorded latencies, they remained silent about developmental changes in children's strategy *efficiencies*. Later studies, combining latencies and strategy reports showed that, with increasing age, children show changes in both strategy selection (e.g., a transition from nonretrieval strategies to retrieval strategies) and strategy efficiency (i.e., strategies becomes faster and less error-prone). This has been shown for addition (e.g., Bisanz et al., 1995; Geary, 1996; Geary et al., 1991; Geary, Bow-Thomas, Liu, & Siegler, 1996a; Siegler & Jenkins, 1989), subtraction (e.g., Robinson, 2001), and multiplication (e.g., Cooney et al., 1988; Lemaire & Siegler, 1995; Mabbott & Bisanz, 2003; Steel & Funnel, 2001). For division, age-related increases have been shown for strategy efficiency but not for retrieval frequency (Robinson et al., 2006).

Importantly, children do not simply use a particular strategy until a better one comes along. Instead, they have several strategies available to them at any time. It is the frequency of use for each strategy that changes across development (cf. Siegler's overlapping waves model, 1995, 1996). Retrieval and nonretrieval strategies might thus co-exist for a period of time in a child's development (Siegler & Jenkins, 1989; Siegler & Shrager, 1984). However, with growing age and more experience in solving arithmetic problems, children's strategy choices become increasingly adaptive, in the sense that the strategy choices are calibrated increasingly precisely to the characteristics of the problems (Lemaire & Siegler, 1995).

THE ROLE OF WORKING MEMORY IN CHILDREN'S ARITHMETIC PERFORMANCE

The role of working memory in children's arithmetic strategy use is investigated in both developmental studies reported in this thesis (cf.

Chapters 6 and 7). As both arithmetic performance (see above) and working-memory capacity (Kail, 1990) increase through the elementary school years, it is not unthinkable that working memory might play a role in children's arithmetic performance. Moreover, the basic modular structure of working memory (with an executive, a phonological, and a visuo-spatial component) has been shown to fit the data of children from 6 years of age (e.g., Gathercole, Pickering, Ambridge, & Wearing, 2004; Swanson, 2006). In the current dissertation, we focused on the role of executive working-memory components in the development of arithmetic performance. The possible role of phonological and visuo-spatial working-memory components in children's arithmetic development is postponed to the general discussion (cf. Chapter 8).

In a series of correlation studies, Bull and colleagues provided extensive evidence for the important role of working memory in children's arithmetic problem solving (e.g., Bull & Johnston, 1997; Bull, Johnston, & Roy, 1999; Bull & Scerif, 2001). Later studies confirmed that executive working-memory processes influence the age-related improvements in arithmetic problem solving, beyond what other processes – such as short-term memory, long-term memory, and processing speed – may contribute (e.g., Swanson, 2004; Swanson & Beebe-Frankenberger, 2004). The great role of working memory in the development of arithmetic skill was also confirmed by Rasmussen and Bisanz (2005), who observed that measures of working memory accounted for a substantial proportion of the variability in arithmetic performance in both preschool children ($R^2 \geq .40$) and 1st grade children ($R^2 \geq .42$). More recently, Swanson and Kim (in press) showed that children's mathematical performance was predicted by executive working memory, phonological short-term memory, and processing speed. However, executive working memory was the best predictor, above the contribution of phonological short-term memory and processing speed. These three capacities jointly accounted for over 74% of children's mathematical performance. Finally, Passolunghi, Vercelloni, and Schadee (in press) performed a longitudinal study that investigated the precursors of mathematics learning. Early mathematics learning was best predicted by working memory and counting ability. When these two variables were included, neither phonological awareness nor intelligence directly influenced children's mathematics abilities. Passolunghi and colleagues concluded that working memory – but not short-term memory – is a distinct and significant predictor of mathematics learning at the beginning of primary school.

Although very interesting, none of the studies mentioned above (a) investigated the role of working memory online, that is to say, with the selective interference paradigm, or (b) investigated the relationship between working memory, on the one hand, and arithmetic strategy selection and strategy efficiency, on the other. In the developmental studies presented in the current thesis, both strategy selection and strategy efficiency were investigated. To that end, we applied the combined approach by obtaining strategy reports, response latencies, and accuracies. In Chapter 6, the role of working memory was tested by means of the individual-difference approach (i.e., by measuring children's working-memory span), whereas in Chapter 7, the role of working memory was investigated by means of the selective interference paradigm (i.e., by imposing a secondary task).

INDIVIDUAL DIFFERENCES

Mathematical ability has been recognized as an important dimension of human intelligence for some time (Spearman, 1927; Thurstone, 1938). Hence, individual differences in mathematical cognition have been studied extensively by cognitive psychologists during recent decades (e.g., Geary, 1993, 2003) and their relationship to online arithmetic performance has been labeled as 'a priority for the field' (LeFevre et al., 1996a). Indeed, it is important to consider the possibility that cognitive processes might differ as a function of individual-difference variables. One of the earliest (and best-known) individual-difference studies has been conducted by Siegler (1988a; see also Kerkman & Siegler, 1997), in which three types of children were revealed: good students (i.e., frequent retriever users), not-so-good students (i.e., infrequent retrieval users), and perfectionists (i.e., children who are good retrievers but rather choose to use nonretrieval strategies). Very recently, Hecht (in press) revealed the same three types in adults. In the current doctoral dissertation, a somewhat different approach was used. More specifically, we tested whether individual-difference variables – such as gender, math anxiety, or working-memory span – were related to adults' simple-arithmetic performance (cf. Chapters 2 and 4). As the study of arithmetic development involves not only documenting age-related changes in strategy selection and strategy efficiency, but also identifying changes in the variables that affect these aspects of performance, the influence of children's individual-difference variables was tested in Chapter 7. In the following paragraphs, an overview of previous individual-difference studies is provided.

ARITHMETIC SKILL

This ability (also called mathematical fluency) refers to a general account of a person's arithmetic abilities, and should be viewed as a continuum from 'novice' to 'expert'. It is generally tested with the French kit (French, Ekstrom, & Price, 1963), a timed pen-and-paper test composed of arithmetic problems of increasing difficulty. Generally, high-skill participants (both children and adults) are faster and less erroneous than low-skill participants in solving arithmetic problems (e.g., Campbell & Xue, 2001; Geary et al., 1992; Gilles, Masse, & Lemaire, 2001; Hecht, Torgesen, Wagner, & Rashotte, 2001; Kaye et al., 1989; Smith-Chant & LeFevre, 2003; Widaman & Little, 1992; Widaman et al., 1992). Skill-related differences have also been found in other numerical tasks, such as counting series (e.g., "5-10-15-20"; LeFevre & Bisanz, 1986), number matching (e.g., LeFevre et al., 1991), arithmetic word problem solving (e.g., Kail & Hall, 1999), and computational estimation (e.g., Levine, 1982). These results were interpreted as evidence for individual differences in calculation processes and individual differences in the obligatory activation of arithmetic facts in long-term memory.

More relevant to the current thesis, however, are studies in which individual differences in arithmetic skill have been related to individual differences in arithmetic *strategy* use. For instance, a significant correlation between arithmetic skill and strategy selection – with high-skill participants retrieving more frequently than low-skill participants – has repeatedly been reported (e.g., Bull & Johnston, 1997; Bull et al., 1999; Geary & Wiley, 1991; Geary et al., 1992; LeFevre et al., 1996a, 1996b; Smith-Chant & LeFevre, 2003; Thevenot, Fanget, & Fayol, in press; Torbeyns et al., 2002, 2004a, 2004b; but see Kirk & Ashcraft, 2001). In the current doctoral dissertation, the relation between arithmetic skill and simple-arithmetic performance was examined more thoroughly by fractionating simple-arithmetic performance in three components (i.e., strategy selection, retrieval efficiency, and nonretrieval efficiency).

MATH EXPERIENCE

In the first empirical chapter of this thesis (cf. Chapter 2), the role of daily arithmetic practice was investigated. Because we observed that this individual-difference variable had significant effects on people's strategy efficiency and strategy selection, we decided to investigate the role of this variable further. Hence, math experience was included in Chapter 5 as well. Note that we are unable to provide a literature overview concerning the role of this variable, as we were the first to test possible effects of math experience on people's strategic performances in mental arithmetic. If effects of math experience would be found, this variable should not be overlooked in future research.

SHORT-TERM MEMORY AND WORKING MEMORY

Short-term memory refers to the passive storage of information, whereas working memory occupies both storage and processing of information. Working memory thus provides attentional processes to keep the short-term memory contents in an activated state. The differentiation between short-term memory and working memory can be made in children (e.g., Kail & Hall, 2001; Swanson & Kim, in press; but see Hutton & Towse, 2001) and in adults (e.g., Engle, Tuholski, Laughlin, & Conway, 1999b; Heitz, Unsworth, & Engle, 2005). The tasks used to test the capacity of both memory systems differ: simple-span tasks are used to test short-term memory capacity (e.g., digit span, letter span, word span), whereas complex-span tasks are used to test working-memory capacity (e.g., reading span, listening span, operation span). Generally, performance on cognitive tasks is better predicted by working-memory span than by short-term memory span (e.g., Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, 2002; Engle et al., 1999b; Just & Carpenter, 1992; Kyllonen & Christal, 1990; Miyake, 2001).

As it has been shown that short-term memory and working-memory capacities grow with age (e.g., Case, 1985; Fry & Hale, 1996, 2000), these variables have mainly been investigated in their relation with development. Relations between short-term memory capacity or working-memory capacity, on the one hand, and arithmetic performance, on the other, have frequently been shown in mathematically disabled children (e.g., Geary et al., 1991, 1999, 2000a; Hitch & McAuley, 1991; McLean & Hitch, 1999; Passolunghi & Siegel, 2001, 2004; Siegel & Ryan, 1989; Swanson, 1993; Swanson & Beebe-Frankenberger, 2004; Swanson & Sachse-Lee, 2001; Turner & Engle, 1989; but see Temple & Sherwood, 2002). Correlations between working-memory capacity and arithmetic performance have been observed in gifted and precocious children as well (e.g., Dark & Benbow, 1991; Geary & Brown, 1991; Swanson, 2006). The relation between short-term memory or working-memory capacities and normally developing children's arithmetic performance has been investigated only recently (e.g., Hecht et al., 2003; Hitch, Towse, & Hutton, 2001; Lépine, Barrouillet, & Camos, 2005; Mabbott & Bisanz, 2003; Noël et al., 2004; Rasmussen & Bisanz, 2005; Swanson, 2004; Swanson & Kim, in press; van der Sluis, de Jong, van der Leij, in press). Swanson (2004), for example, showed that short-term memory and working memory operated independently of each other, and that only working memory contributed unique variance to children's arithmetic performance. These results were confirmed by Swanson and Kim (in press), who showed that working memory predicted children's mathematical performance independent of contributions of short-term memory and processing speed.

In our developmental studies, we tested the role of working-memory capacity (Chapter 6) and short-term memory capacity (Chapter 7) in normally-developing children. The inclusion of strategy reports in both studies enabled us to verify whether short-term memory and working-memory capacities were related to strategy selection, strategy efficiency, or both. These issues are especially relevant as there is an ongoing debate about whether or not working memory is needed in strategy selection. Geary and colleagues (1993a, 1996a, 2004) suggested that children's working-memory capacities are correlated with the frequency of counting but not with the frequency of retrieval. Similarly, Noël et al. (2004) failed to observe a relation between children's working-memory span and their percentage retrieval use. Finally, Barrouillet and Lépine (2005) did observe a relation between children's working-memory span and their percentage retrieval use, albeit only for small problems.

PROCESSING SPEED

The speed with which various information processes occur is another important variable, especially because its relation with developmental courses. Indeed, it has been shown that older children think faster than do younger children (e.g., Fry & Hale, 1996, 2000; Hale, 1990; Kail, 1988, 1991), and that older adults are slower than younger adults (e.g., Salthouse, 1991, 1992, 1996). Hence, the relation between processing speed and arithmetic performance has been studied in developmental studies comparing older with younger children (e.g., Adams & Hitch, 1997; Swanson, 2004) or comparing older with younger adults (e.g., Duverne & Lemaire, 2004). Several tests to measure processing speed, such as visual number matching tasks and cross-out tasks (e.g., Bull & Johnston, 1997), have been used.

It is important to note that processing speed, on the one hand, and short-term memory capacity and working-memory capacity, on the other, are closely related (e.g., Adams & Hitch, 1997; Case, Kurland, & Goldberg, 1982). Indeed, if mental operations are executed faster, more working-memory space is left free for other storage and/or processing tasks. Consequently, age-related increases in processing speed might mediate developmental increases in working-memory capacity (e.g., Fry & Hale, 1996, 2000; Kail, 1992; Kail & Park, 1994). However, processing speed may still be linked to cognitive abilities even when working-memory capacity is statistically controlled for, and vice versa. Whereas some authors observed that processing speed was more predictive of arithmetic performance than short-term memory or working memory were (e.g., Bull & Johnston, 1997; Fuchs et al., 2006; Kail & Hall, 1999), other researchers found exactly the opposite (e.g., Hitch et al., 2001; Noël et al., 2004; Swanson & Beebe-Frankenberger, 2004; Swanson & Kim, in press). In the current thesis, a measure of processing speed has been included in one of our developmental studies (cf. Chapter 7). As working memory was tested as well, we were able to disentangle the role of both variables (i.e., processing speed and working memory).

MATH ANXIETY

Math anxiety can be defined as 'feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations' (Richardson & Suinn, 1972, p. 551). The most difficult point in investigating math anxiety is that it correlates with many other variables, such as math competence, test anxiety (Hembree, 1990), gender (e.g., Ashcraft, 1995; Ashcraft & Faust, 1994; Hyde, Fennema, Ryan, Frost & Hopp, 1990), working-memory capacity (Ashcraft & Kirk, 2001), and career choice (e.g., Chipman, Krantz, & Silver, 1992). However, there is no complete confounding between these variables (e.g., Faust, Ashcraft, & Fleck, 1996; Hembree, 1990). Moreover, math anxiety shows only a small relationship to general intelligence ($r = -0.17$; Hembree, 1990), which makes the specific role of math anxiety in mental arithmetic, after all, worth exploring. Math anxiety can be tested with the Mathematics Anxiety Rating Scale (MARS; Richardson & Suinn, 1972) or its abbreviated version, the short MARS (Alexander & Martray, 1989). However, easier procedures can be used as well. Fleck (1994), for example, asked "On a scale from 1 to 5, how math anxious are you?" This item correlated 0.85 with the short MARS. A comparable measure has been used in our studies (cf. Chapters 5 and 7).

Anxiety-related effects are interesting because these emotional effects can be framed within a cognitive theory. Indeed, according to the processing efficiency theory (Eysenck & Calvo, 1992), math anxiety induces intrusive thoughts and worries, which then compete with the ongoing arithmetic task for limited processing resources. Accordingly, math anxiety effects are supposed to be especially apparent in those tasks tapping the limited capacity of working memory (i.e., complex-arithmetic tasks). This hypothesis has been confirmed: High-anxious adults have been shown to perform worse on complex-arithmetic tasks (Ashcraft, 1995; Ashcraft & Faust, 1994; Ashcraft & Kirk, 2001; Faust et al., 1996; Kellogg, Hopko, & Ashcraft, 1999) but not on simple-arithmetic tasks (Ashcraft & Kirk, 1998, 2001; Ashcraft & Faust, 1994; Faust et al., 1996; Kirk & Ashcraft, 2001; Seyler et al., 2003). Accordingly, it has been argued that math anxiety does not affect the basic process of direct fact retrieval (e.g., Ashcraft, 1995). However, the question as whether math anxiety affects strategy selection, strategy efficiency, or both, remains to be answered (but see Faust et al., 1996, who observed no effects of math anxiety on strategy selection). We tried to fill this gap by asking people's math-anxiety level as well as their simple-arithmetic strategy choices. Finding effects of math anxiety would emphasize the – often underestimated – interaction between cognition and emotion.

CALCULATOR USE

Nowadays, when confronted with an arithmetic problem, people might rather grasp a calculator (or another device such as a mobile phone) instead of mentally calculating the response. In fact, technology has somewhat obviated the need to calculate mentally. In that view, we thought it could be interesting to investigate the effects of people's calculator use on their mental-arithmetic performance. This question has been posed earlier, more specifically in cross-cultural studies. For example, Chinese students report less frequent calculator use than do Canadian students (Campbell & Xue, 2001; LeFevre & Liu, 1997). Though, Campbell and Xue (2001) observed no correlations between calculator use and simple-arithmetic performance. In the present doctoral dissertation, we investigated the relation between young adults' calculator use and their simple-arithmetic strategies.

GENDER

Gender differences in arithmetic performance have been studied in children rather than in adults. A longitudinal study of gender differences in children's mathematical thinking (Fennema, Carpenter, Jacobs, Franke, & Levi, 1998) showed that girls used more concrete strategies (such as counting) whereas boys used more abstract strategies that reflected conceptual understanding. This finding has been confirmed in other studies, showing more *frequent* retrieval use in boys than in girls (Carr & Davis, 2001; Carr & Jessup, 1997; Carr, Jessup, & Fuller, 1999; Davis & Carr, 2002). Whether or not boys are more *efficient* retrieval users than girls remains a debated topic. Carr and Davis (2001) observed that boys' retrieval use was more accurate but not faster than that of girls, whereas Royer et al. (1999a) provided preliminal evidence that boys' retrieval use was faster than that of girls. It should be noted, though, that Royer and colleagues' latency analyses included both correctly and incorrectly solved trials. Moreover, they did not obtain strategy reports; the boys' faster simple-arithmetic performance could thus also be due to more frequent retrieval use and/or more efficient nonretrieval use. One study even reported less efficient retrieval use in boys than in girls (Bisanz et al., 1995). Gender differences in the efficiency of children's *non*retrieval use have not yet been observed (Carr & Davis, 2001; LeFevre et al., 2006; but see Biszanz et al., 1995).

In adults, males have been shown to outperform females in mathematical thinking (Mulhern & Wylie, 2004) and in arithmetic problem solving (Geary, Saults, Liu, & Hoard, 2000b). Further, in a meta-analysis of 100 studies investigating gender differences in mathematics, males were found to outperform females in mathematics performance (Hyde, Fennema, & Lamon, 1990). In the current thesis, gender differences were investigated in adults (cf. Chapters 2 and 5) as well as in children (cf. Chapter 7).

STRUCTURE OF THE DOCTORAL DISSERTATION

The goal of the present PhD project was to extend our knowledge on strategies used to solve simple-arithmetic problems. Besides the Introduction (Chapter 1) and the General Discussion (Chapter 8), there are three parts in this thesis. Part I deals with practice-related effects (Chapters 2 and 3), Part II deals with the role of working memory (Chapters 4 and 5), and Part III is concerned with developmental trajectories (Chapters 6 and 7). Each part entails two empirical chapters; and each chapter was written as an individual paper. Hence, the text of some of the chapters may partially overlap. References are provided at the end of this dissertation.

The current thesis also includes two addenda that did not reach publication but do provide an additional and relevant view on the data – especially on the statistical level. In Chapter 2, the addendum contains ex-Gaussian analyses of variance. Such ex-Gaussian analyses are interesting since they provide detailed information about retrieval and nonretrieval efficiencies without relying on verbal strategy reports (Penner-Wilger, Leth-Steensen, & LeFevre, 2002). In Chapter 6, the addendum contains state trace analyses. As faster latencies implicate smaller effect sizes (e.g., Verhaeghen & Cerella, 2002), analyzing developmental data should control for general speeding differences. State trace analyses enabled us to account for such scaling effects.

CHAPTER 2
THE INFLUENCE OF PROBLEM FEATURES AND INDIVIDUAL DIFFERENCES ON STRATEGIC PERFORMANCE IN SIMPLE ARITHMETIC

Memory & Cognition (in press)[1,2]

The present study examined the influence of features differing across problems (problem size and operation) and differing across individuals (daily arithmetic practice, the amount of calculator use, arithmetic skill, and gender) on simple-arithmetic performance. Regression analyses were used to investigate the role of these variables in both strategy selection and strategy efficiency. Results showed that more-skilled and highly practiced students used memory retrieval more often and executed their strategies more efficiently than did less-skilled and less practiced students. Furthermore, calculator use was correlated with retrieval efficiency and procedural efficiency but not with strategy selection. Only very small associations with gender were observed, with boys retrieving slightly faster than girls. Implications of the present findings for views on models of mental arithmetic are discussed.

---

[1] This paper was co-authored by André Vandierendonck and Yves Rosseel.

[2] Thanks are extended to the secondary school 'Immaculata Instituut' in De Panne (Belgium), where all experiments were administered, and to David Geary and two anonymous reviewers for their helpful comments on previous drafts of this article.

INTRODUCTION

Strategic performance of adults consists of two main components. If people want to solve a cognitive problem, they will first have to *choose* the most appropriate strategy to solve it (i.e., strategy selection). Subsequently, they will have to *execute* the chosen strategy with reasonable speed and accuracy (i.e., strategy efficiency). For a long time, mental arithmetic research assumed that adults used only memory retrieval to solve simple-arithmetic problems such as 8 + 3 or 5 x 4 (e.g., Ashcraft, 1987, 1992, 1995; Campbell, 1987a, 1995; Campbell & Oliphant, 1992; Lebiere & Anderson, 1998; McCloskey, 1992; Siegler, 1989; Widaman & Little, 1992). Fairly recently however, LeFevre and colleagues (LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b; see also Baroody, 1994; Geary, Frensch, & Wiley, 1993b; Geary & Wiley, 1991) showed that even skilled adults still make substantial use of procedures such as counting (e.g., 6 + 3 = 6 + 1 + 1 +1) and transformation (e.g., 7 + 5 = 7 + 3 + 2) when solving simple-arithmetic problems. It is clear that retrieval and nonretrieval (i.e., procedural) strategies differ in their efficiency, as retrieval is generally much faster (i.e., more efficient) than any procedural strategy. Although people can still use other strategies to solve arithmetic problems (e.g., using a calculator), the present study investigates mental arithmetic and thus focuses on the two broad kinds of strategy mentioned above: retrieval strategies and procedural strategies.

Both strategy selection and strategy efficiency may depend on several factors such as problem features (e.g., operation, problem size) and individual differences[3] (e.g., arithmetic skill and arithmetic practice).

[3] In the present paper, the term 'individual differences' is used to refer to differences inherent in individuals (e.g., gender) as well as to differences resulting from the environment (e.g., arithmetic practice, calculator use).

Although models have been proposed in which such experiential factors are the main determinants of mental representation, acquisition, and performance (e.g., Ashcraft, 1987; Campbell & Graham, 1985; Siegler, 1988b; Siegler & Shipley, 1995), there are very few direct comparisons of the simple-arithmetic performance of adults differing in mathematical education, arithmetic skill, or arithmetic practice (see also LeFevre and Liu, 1997). Moreover, up until now, no study investigated the effects of these factors in strategy selection and strategy efficiency separately. The present study therefore examined the effects of features differing across problems and across individuals on strategic performance in simple arithmetic.

PROBLEM FEATURES

Although adults rely on both retrieval and procedural strategies for the entire domain of elementary arithmetic (i.e., the four basic operations; e.g., Campbell and Xue, 2001), they adjust their strategy selection to the *operation* that has to be applied. Adults' solving of subtractions and divisions for example, relies more heavily on procedural strategies than either addition or multiplication, for which retrieval strategies are predominantly used (Campbell & Xue, 2001; Seyler, Kirk, & Ashcraft, 2003). Furthermore, multiplications are solved even more frequently by means of direct retrieval than additions (e.g., Hecht, 1999; LeFevre et al., 1996a, 1996b). This might be explained by the fact that multiplications are for the most part based on declarative knowledge whereas additions would be based on both declarative and procedural knowledge (Roussel, Fayol, & Barrouillet, 2002). Whatever the operation is, strategy selection also depends on the *problem size*. For smaller problems (e.g., 2 + 4), people generally retrieve answers from their long-term memory, but for larger problems (e.g., 8 + 6) they are more inclined to use procedural strategies. Because procedural strategies are less efficient than retrieval, longer latencies and higher error rates are observed on larger problems. This very robust effect is known as the problem-size effect, which indicates that solution times and error percentages increase as problem size increases. More frequent retrieval use on small problems than on large problems is not the only source for the problem-size effect, though. According to Campbell and Xue (2001), there are as many as three sources for the problem-size effect: more frequent use of procedural strategies for large than for small problems, a lower efficiency of retrieval strategies for large than for small problems, and a lower efficiency of procedural strategies for large than for small problems. Lower retrieval efficiencies can be explained by weaker associative connections between problem-answer pairs in the retrieval network (e.g., Siegler, 1988b), whereas lower procedural efficiencies can be explained by the larger number of sub-operations to be performed. Nevertheless, the problem-size effect is not only based on strategic sources; the structure of the mental network (with different network strengths and spreading activation characteristics), and interference from competing associations are other contributing factors (e.g., Zbrodoff, 1995).

INDIVIDUAL DIFFERENCES

Strategy selection and strategy efficiency do not only depend on problem features, but also on individual traits and culture-based factors. Several studies indeed found differences between East-Asians' and North-Americans' simple-arithmetic performance (e.g., Chen & Uttal, 1988; Campbell & Xue, 2001; Geary, 1996; Geary, Bow-Thomas, Liu, & Siegler, 1996a; Geary, Fan, & Bow-Thomas, 1992; Geary, Salthouse, Chen, & Fan, 1996b; Geary et al., 1997; LeFevre & Liu, 1997; Penner-Wilger, Leth-Steensen, & LeFevre, 2002; Stevenson, Chen, & Lee, 1993; Stevenson, Lee, & Stigler, 1986; Stevenson et al., 1990; reviewed by Geary, 1994). Because North-Americans frequently use procedures whereas East-Asians rely primarily on memory retrieval, faster and less error-prone arithmetic performance is observed in the latter group than in the former group. Arithmetic performance differences are not only found across cultures though, but also within one single culture. Persons may differ from each other in several respects such as mathematical education, daily arithmetic practice, arithmetic skill, et cetera. However, there has been little research examining such differences in adults' cognitive processes in arithmetic. In the following, we summarize the main results of studies examining effects of cognitive factors on arithmetic performance. We also consider the role of gender.

Differences in arithmetic performance have been found to depend on *arithmetic skill*. LeFevre and Bisanz (1986) for example, found that low-skill persons used less efficient and slower mental calculation processes than did high-skill persons. Therefore, the difference between low-skill and high-skill participants was larger on items that required calculations than on items that could be solved without calculations. Furthermore, LeFevre et al. (1996a, 1996b) observed more frequent retrieval use in high-skill than in low-skill persons. More recently, high-skill participants were shown to be more efficient (i.e., faster) in solving simple-arithmetic problems than low-skill participants (Campbell & Xue, 2001; Kirk and Ashcraft, 2001, see also Geary & Widaman, 1987; Gilles, Masse, & Lemaire, 2001). Finally, LeFevre and colleagues (LeFevre & Kulak, 1994; LeFevre, Kulak, & Bisanz, 1991) found evidence for individual differences in the obligatory activation of addition facts. As associative connections are stronger in high-skill participants than in low-skill participants, they concluded that accessibility of arithmetic facts may contribute to individual differences in the solution of arithmetic problems.

Besides arithmetic skill, strategic performance may also depend on other factors such as math attainment, daily arithmetic practice, and calculator use. Hecht (1999), for example, showed that adults with higher levels of *math attainment* used retrieval strategies more frequently, were more accurate in solving math facts, and retrieved arithmetic problems faster than did adults with lower levels of math achievement. Roussel et al. (2002) found that people with high amounts of *daily arithmetic practice* (primary school teachers) exhibited smaller problem-size effects than people with low amounts of daily arithmetic practice (undergraduate psychology students). However, the highly practiced participants were found not to differ from the less practiced participants in the strategies they used (i.e., strategy selection). The frequency of *calculator use* might be another influencing factor. From primary school on, children are taught how to use a hand-held calculator. However, calculators themselves are at the centre of several controversies, not only educational (i.e., is it good for children to use calculators?) but also conceptual (i.e., are calculators designed and implemented well?, Thimbleby, 2000). Very few studies have investigated effects of calculator use on simple-arithmetic performance though. Campbell and Xue (2001) observed no reliable effect of the frequency of calculator use on simple-arithmetic strategy selection or strategy efficiency.

Gender differences have been found in young children's arithmetic strategy selection (e.g., Carr & Jessup, 1997; Carr, Jessup, & Fuller, 1999; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998). More specifically, girls are more likely than boys to use procedural strategies whereas boys are more likely than girls to use retrieval strategies. These gender differences in simple-arithmetic strategy selection have been shown to be driven not only by skill differences, but also by girls' and boys' strategy preferences (Carr & Davis, 2001). Furthermore, gender differences have been found in retrieval efficiency as well, with boys being faster than girls from fifth grade on (Royer, Tronsky, Chan, Jackson, & Marchant, 1999a). Carr and Davis (2001), in contrast, observed no differences between boys and girls in retrieval efficiency. Although no study explicitly investigated whether gender differences in strategy selection and strategy efficiency exist in adults, Geary, Saults, Liu and Hoard (2000b) re-analyzed simple-arithmetic performance data obtained in an earlier study by Geary et al. (1993b). They observed a trend to more frequent retrieval use by men than by women (86% vs. 66%, $p < .07$) but no gender differences in retrieval efficiency.

OVERVIEW OF THE PRESENT STUDY

As described earlier, problem features have large effects on adults' simple-arithmetic performance. Individual differences have less frequently been studied, though. For example, gender effects in adults' simple-arithmetic performance have not been investigated thus far. Effects of the frequency of calculator use have not been observed either. The effects of arithmetic practice over and above those of arithmetic skill are still debated as well. Moreover, even if individual differences in simple-arithmetic performance were studied, no distinction was made between their role in strategy selection and strategy efficiency. Therefore, the present study investigated effects of both problem features and individual differences on simple-arithmetic strategy selection and strategy efficiency separately. In addition to the investigation of two typical problem features (problem size and operation), effects of four individual traits were tested: daily arithmetic practice, arithmetic skill, the amount of calculator use, and gender. The novelty here is the distinction between skill and practice. Whereas most previous studies selected low-skill and high-skill participants based on arithmetic subtests only (e.g., Gilles et al., 2001; Hecht, 1999; LeFevre & Bisanz, 1986, LeFevre et al., 1996b), the present study incorporated the amount of daily arithmetic practice in addition to the measure of arithmetic skill. In fact, the operationalization of 'daily arithmetic practice' was based on the students' high-school curricula, which all differed in the number of mathematic and scientific hours per week. The Belgian education system, in which high-school students have to choose one main class every two years, offers a good opportunity to investigate such practice effects in an ecologically valid way. Indeed, all students in the present study were enrolled in a specific curriculum with a fixed amount of arithmetic hours each week. In contrast, arithmetic skill was measured by means of a frequently used pen-and-paper test (the French kit). A short questionnaire determined each student's habits concerning calculator use.

Based on previous research, we expected more frequent retrieval use and more efficient strategy use for the high-skill participants compared to the low-skill participants. We also expected simple-arithmetic performance to be related to daily arithmetic practice, with more frequent retrieval use and more efficient strategy execution in the more-practiced students than in the less-practiced students. Although there is no evidence for effects of calculator use on simple-arithmetic performance (Campbell & Xue, 2001), we expected less frequent retrieval use and lower strategy efficiencies for participants frequently using the calculator than for participants rarely using the calculator. Concerning gender differences, it was hard to make any predictions, because such differences are more pronounced in children than in adults. However, on the basis of these developmental studies, if any differences were to appear in the present study, we expected more frequent and more efficient retrieval use for male than for female students.

METHOD

PARTICIPANTS

Sixty secondary-school students of the sixth year participated in this study[4]. The amount of daily arithmetic practice variable was operationalized by the number of hours per week dedicated to arithmetic and scientific

[4] In Belgium, secondary-school education starts at 12 years of age and normally takes six years to finish. After their second year of secondary school, students have to choose among different study options, such as Humanities, Economics, Languages, Mathematics, or Sciences. The amount of daily arithmetic practice (defined as the number of hours per week dedicated to arithmetic and scientific classes) varies across these study options. More specifically, students enrolled in a Mathematics curriculum have more arithmetic-related hours than students in an Economics curriculum, which have in their turn more arithmetic-related hours than students in a Humanities curriculum.

classes. Scientific classes were considered because they are classes in which arithmetic is frequently used. The number of arithmetic and scientific hours per week was 3 (15 students, mean age: 17 years 8 months, 14 girls and 1 boy), 4 (16 students, mean age: 17 years 7 months, 14 girls and 2 boys), 9 (2 students, mean age: 17 years 6 months, 1 girl and 1 boy), 11 (8 students, 4 girls and 4 boys, 17 years 9 months), 13 (10 students, mean age: 17 years 6 months, 1 girl and 9 boys), or 15 (9 students, mean age: 17 years 8 months, 5 girls and 4 boys). At the time of measurement, all students were enrolled in their specific curriculum for at least one and a half year. All students participated voluntarily, with permission of their parents and the school teachers.

PROCEDURE AND STIMULI

Each participant was tested individually in a quiet class room for approximately 45 minutes. The test session was started with short questions about the participant's age, study curriculum, and the number of arithmetic and scientific hours per week. Three tasks were given to each participant. The first one was the simple-arithmetic task, which consisted of two blocks, one with addition problems and one with multiplication problems, the order of which was counterbalanced across participants. Subsequently, an arithmetic skill test (the French kit) was administered. The session ended with a short questionnaire about calculator use. In the following, each task is described more in detail.

Simple-arithmetic task. Stimuli of the simple-arithmetic task consisted of simple additions and simple multiplications. As in previous research, we used the so-called standard set of problems (LeFevre et al., 1996b), which excludes problems involving 0 or 1 as an operand or answer. Both addition and multiplication problems were composed of pairs of numbers between 2 and 9, with tie problems (e.g., 3 + 3) excluded. Because commuted pairs (e.g., 2 + 4 and 4 + 2) were considered as two different problems, this resulted in 56 addition problems (ranging from 2 + 3 to 8 + 9) and 56 multiplication problems (ranging from 2 x 3 to 8 x 9). Problem size was defined as the correct answer to the problem (from 5 to 17 for the sums and from 6 to 72 for the products). This continuous definition of problem size differs from the widely used dichotomous definition of problem size (i.e., a categorization into small and large problems).

A trial started with a fixation point, which appeared for 500 msec. Then the arithmetic problem appeared in the center of the screen. The addition and multiplication problems were presented horizontally in Arabic format as dark-blue characters on a light-grey background, with the operation sign (+ or x) at the fixation point. The problem remained on the screen until the subject responded. Although participants were required to respond as quickly and as accurately as possible, no time deadline was set, because it has been shown that a fast deadline increased reported use of retrieval, especially for large problems (Campbell & Austin, 2002). Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, participants wore a microphone, which was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 msec. On each trial, feedback was presented to the participants, a green 'Correct' when their answers were correct, and a red 'Fout' (i.e., Dutch for 'Wrong') when their answers were wrong.

Participants were also told to report the strategy they used for each single problem. They could choose one of the following strategy categories: 'Retrieval', 'Counting', 'Transformation', and 'Other' (see e.g., Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler et al., 2003). At the beginning of the experiment, each strategy was described as follows: *(1) Retrieval: You solve the problem by just remembering or knowing the answer directly from memory. It means that you know the answer without any additional processing, or that the answer "pops into your head", (2) Counting: You solve the problem by* *counting a certain number of times to get the answer (e.g., 6 + 3 = 6… 7… 8… 9; 3 x 6 = 6… 12… 18), (3) Transformation: You solve the problem by referring to related operations or by deriving the answer from some known facts. You change the presented problem to take advantage of a known arithmetical fact (e.g., 6 + 7 = 6 + 6 + 1; 9 x 6 = 60 – 6), (4) Other: You solve the problem by using a strategy unlisted here (e.g., guessing), you used more than one strategy, or you do not know what strategy you used to solve the problem.* After each trial, the four category names were displayed on the screen. The participant also kept a copy of the strategy report instructions for reference during the study. It was emphasized that the presented strategies were not meant to encourage use of a particular strategy.

The answer of the participant, the reported strategy, and the validity of the trial were recorded online by the experimenter. All invalid trials (e.g., failures of the voice-activated relay) were discarded and rerun at the end of the block. This procedure enabled us to minimize data loss due to unwanted failures. Each block (addition or multiplication problems) started with 4 practice problems, followed by the experimental problems. As each problem was presented twice, each block consisted of 112 arithmetic trials, which were presented in a random order. After the first block and a short break, the second block (with the other operation) was administered, consisting of 4 practice problems and 112 experimental problems as well.

French kit*.* After the simple-arithmetic task, participants completed two arithmetic subtests of the French kit (Ekstrom, French, & Harman, 1976; French, Ekstrom, & Price, 1963), one page of complex addition problems and one page of complex subtraction and multiplication problems. Each page contained six rows of ten vertically oriented problems. Each participant was given two minutes per page to solve the problems as quickly and accurately as possible. Arithmetic skill was defined as the total number of problems solved correctly on both tests. This measure of arithmetic skill reflects the ability to quickly and accurately execute strategies on multi-digit problems. The French Kit is also used to measure arithmetic fluency and working-memory management (e.g., carrying and borrowing; Geary & Widaman, 1992).

Calculator-use questionnaire. Participants received a page on which the following question was written: "How often did you use a calculator (or another electronically device, e.g., cell phone) when doing arithmetic problems (e.g., 65 + 34, 23 x 17)?" Participants had to provide an answer to this question by marking a 5-point rating scale ranging from "never" to "always", once concerning their experiences during elementary school and once concerning their experiences during secondary school.

RESULTS

Overall, 1305 trials (i.e., 9.09%) were spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials spoiled due to failures of the sound-activated relay to 425 (i.e., 2.96%). Further, all incorrect trials and all trials on which participants reported having used a strategy 'Other' were deleted (i.e., 3.03%). Finally, all the response times (RTs) more than 4 standard deviations from the participant's mean (per operation) were discarded as outliers (0.5% for addition and 1.1% for multiplication). The final data set consisted of 13026 valid trials, which corresponds to a total data loss of less than 8%. In the following, all reported results are considered to be significant if $p < .05$, unless mentioned otherwise.

Regression analyses were performed to detect which factors contributed to strategy selection and strategy efficiency. Regression analyses were run for additions and multiplications separately on the three dependent variables: percentages retrieval use, retrieval RTs, and procedural RTs (see Table 1). Predictors in all regression analyses were: (1) problem size (defined as the correct answer to the arithmetic problem), (2) daily arithmetic practice (defined as the number of mathematic and scientific hours per week), (3) calculator use (as measured by the questionnaire), (4) arithmetic skill (i.e., score on the French kit), and (5) gender (male or female). The first predictor varies across problems whereas the other variables vary across individuals.

Table 1
Means, standard deviations (SD), medians, minima, and maxima of the three dependent variables used in the regression analyses.

Addition	Mean	SD	Median	Minimum	Maximum
Retrieval use (%)	78	4	81	31	100
Retrieval RTs (msec)	874	296	857	601	1382
Procedural RTs (msec)	1114	680	1035	710	2464
Multiplication	**Mean**	**SD**	**Median**	**Minimum**	**Maximum**
Retrieval use (%)	87	3	89	53	100
Retrieval RTs (msec)	1376	688	1343	689	2573
Procedural RTs (msec)	2865	2368	2184	889	7453

Table 2 describes means, medians, minima and maxima of the independent variables varying across individuals (except gender)[5]. A paired-samples *t*-test showed that calculator use was more frequent in secondary school than in elementary school, $t(59) = 13.40$, $SE = .157$, with mean scores of 4.0 ($SD = .88$) and 1.9 ($SD = .90$), respectively. As we expected that only the current frequency of calculator use would influence strategic performance, the frequency of calculator use in secondary school was included in the regression analyses whereas the frequency of calculator use in elementary school was not. A summary of all the regression analyses is presented in Table 3.

The R^2 for percentage *retrieval use* was .298, $F(5,774) = 65.72$ for addition ($MSe = 764.08$) and .164, $F(5,1493) = 58.67$ for multiplication ($MSe = 639.55$). For both operations, retrieval use occurred more frequently with smaller problem sizes and with higher arithmetic skill. More frequent daily practice predicted more frequent retrieval use only for multiplications. The regression analyses on *retrieval RTs* resulted in an R^2 of .322, $F(5,722) = 68.51$ for addition ($MSe = 48515.37$) and an R^2 of .248, $F(5,1417) = 93.51$ for multiplication ($MSe = 307053.80$). For both operations, answers were retrieved faster with smaller problem size, higher arithmetic skill, and less frequent calculator use. More extensive daily practice predicted faster retrieval use only for multiplications. Furthermore, boys tended to be slightly faster in retrieving multiplication facts than girls. Finally, for *procedural RTs* an R^2 of .175, $F(5,328) = 13.87$ was obtained for addition (MSe =

[5] It should be noted that some predictor variables correlated with each other. More specifically, the correlation between daily arithmetic practice and arithmetic skill, between calculator use and arithmetic skill, and between daily arithmetic practice and calculator use were significant ($r = .256$, $r = -.347$, and $r = -.297$, respectively). These correlations are not problematic for the regression analyses, however. Indeed, omitting one of these predictors in the regression model did not result in dramatic changes of the parameter estimates or significance results for the remaining predictors.

240580.01) and an R^2 of .229, $F(5,350) = 20.77$ for multiplication (MSe = 4013325.39). For both operations, procedural strategies were performed faster when problem size was smaller, when arithmetic skill was higher, and when calculator use was less frequent. Once more, high daily practice predicted faster procedural use only for multiplication but not for addition.

Table 2
Means, medians, minima, and maxima of the individual-characteristic variables (except gender) used as predictors in the regression analyses.

	Mean	Median	Minimum	Maximum
Daily arithmetic practice (hours)	8	4	3	15
Calculator use secondary school [a]	4	4	1	5
Calculator use elementary school [a]	2	2	1	5
Arithmetic skill (score on the French kit)	28	26	15	58

[a] The frequency of calculator use was questioned for both the years at the elementary school and the years in secondary school. Only the current frequency of calculator use (i.e., in secondary school) was used in the regression analyses, however.

Table 3 (continued on next page)
Summary of the regression analyses for variables predicting percentage retrieval use, retrieval RTs, and procedural RTs.

Retrieval use	Addition		
	B	*SE B*	*ß*
Problem size	-4.348	.265	-.495**
Arithmetic skill	.697	.131	.174**
Calculator use	-.099	.052	-.066
Daily practice	.044	.244	.006
Gender	3.061	2.614	.044
Retrieval use	**Multiplication**		
	B	*SE B*	*ß*
Problem size	-.595	.037	-.380**
Arithmetic skill	.367	.087	.109**
Calculator use	.065	.034	.052
Daily practice	.422	.161	.074**
Gender	.732	1.725	.013

* $p < 0.05$. ** $p < 0.01$.

Table 3 (continued on next page)
Summary of the regression analyses for variables predicting percentage retrieval use, retrieval RTs, and procedural RTs.

Retrieval RTs	Addition		
	B	*SE B*	*ß*
Problem size	30.873	2.231	.424**
Arithmetic skill	-9.206	1.068	-.288**
Calculator use	1.510	.423	.125**
Daily practice	-1.011	2.042	-.018
Gender	-38.779	22.061	-.070
Retrieval RTs	**Multiplication**		
	B	*SE B*	*ß*
Problem size	14.177	.849	.385**
Arithmetic skill	-18.302	1.930	-.238**
Calculator use	2.037	.779	.069**
Daily practice	-7.751	3.625	-.059*
Gender	-80.402	38.564	-.060*

* $p < 0.05$. ** $p < 0.01$.

Table 3
Summary of the regression analyses for variables predicting percentage retrieval use, retrieval RTs, and procedural RTs.

Procedural RTs	Addition		
	B	*SE B*	*ß*
Problem size	40.437	9.270	.220**
Arithmetic skill	-9.072	3.740	-.138**
Calculator use	3.401	1.740	.127*
Daily practice	-13.553	7.366	-.114
Gender	-131.687	76.057	-.109
Procedural RTs	**Multiplication**		
	B	*SE B*	*ß*
Problem size	19.154	6.203	.146**
Arithmetic skill	-108.738	17.571	-.309**
Calculator use	12.721	5.235	.146**
Daily practice	-103.214	28.510	-.214**
Gender	344.256	346.783	.070

* $p < 0.05$. ** $p < 0.01$.

GENERAL DISCUSSION

Results of the present study showed that Belgian high-school students used a variety of strategies to solve simple-arithmetic problems, which is in accordance with comparable research in non-European participants (e.g., Campbell & Xue, 2001; Hecht, 1999; LeFevre & Liu, 1997; LeFevre et al., 1996a, 1996b), and – as LeFevre and collaborators reported – sharply in contrast with the assumption that adults always retrieve arithmetic facts from memory. There are, however, both similarities and differences between the present results and previous findings. We observed 78% retrieval use for addition and 87% retrieval use for multiplication. These percentages are at the high end of the range of percentages observed in North-America, reaching from 66% to 76% for addition and from 59% to 96% for multiplication (e.g., Campbell & Xue, 2001; Campbell & Timm, 2000; Geary, 1996; Hecht, 1999; LeFevre et al., 1996a, 1996b). They are, however, well beneath the percentages observed in East-Asia, with reported percentages of retrieval use of 92% for addition and 100% for multiplication (e.g., Campbell & Xue, 2001; Geary, 1996; LeFevre & Liu, 1997). The number of participants using retrieval on all trials was 6 (10%) for addition and 16 (26%) for multiplication, these figures are comparable to those of LeFevre et al. (1996a, 1996b) with 12.5% and 28% of the participants using retrieval on all trials, for addition and multiplication respectively.

Over and above confirming previous results, the present study yielded several new findings concerning individual differences in strategy selection and strategy efficiency. Concerning *strategy selection*, direct memory retrieval was used more often (a) by high-skill students than by low-skill students, and (b) by more-practiced students than by less-practiced students – the latter being true for multiplication but not for addition. Note that some previous studies did observe individual differences in strategy selection (e.g., Hecht, 1999; LeFevre et al., 1996b), whereas others did not (e.g., Roussel et al., 2002). *Strategy efficiency* also differed across individuals. Both retrieval efficiency and procedural efficiency increased (a) with the level of arithmetic skill and (b) with the level of daily arithmetic practice – the latter being true for multiplication but not for addition. Remarkably, the frequency of calculator use influenced strategy efficiency as well: students frequently using the calculator showed lower retrieval and procedural efficiency levels but did not differ in strategy selection. Gender only correlated with retrieval efficiency, indicating that boys were slightly faster than girls in retrieving multiplication facts from memory. In the following we first elaborate on the observed individual differences and then describe some implications for the present models of mental arithmetic.

DAILY ARITHMETIC PRACTICE

As most previous studies examined individual differences by means of pen-and-paper tests, the present study also incorporated a more 'ecological' variable, namely daily arithmetic practice. In fact, the amount of daily arithmetic practice was based on the number of arithmetic and scientific hours per week during the past years in secondary school. As outlined in the introduction, practice effects were expected on strategy selection, retrieval efficiency, and procedural efficiency. All these hypotheses were confirmed: All three measures of simple-arithmetic strategic performance were significantly linked to arithmetic practice, albeit only for multiplication problems.

The finding that less-practiced students used retrieval less often than more-practiced students is in agreement with previous studies assuming that practice may lead to increases in retrieval frequency (e.g., Siegler, 1988b; Widaman & Little, 1992). Although strategy selection depended on arithmetic practice, both retrieval and procedural strategies were used irrespective of the level of practice. Moreover, more-practiced students used retrieval more efficiently than did less-practiced students. Frequently practiced problems may have developed stronger problem-answer associations than less frequently practiced problems (e.g., Siegler, 1988b), and these stronger problem-answer associations may have resulted in faster retrieval use. Previous studies indeed showed that associative connections may differ across individuals (LeFevre & Kulak, 1994; LeFevre et al., 1991). Procedure efficiency was much higher for the more-practiced students than for the less-practiced students. LeFevre et al. (1996a) stated that practice may lead not only to more frequent retrieval use, but may lead to automatic activation of procedural strategies as well. If the successful use of procedural strategies increases the strength of the problem-procedure association, practice may thus also influence procedural efficiency, an effect that was observed here. More-practiced students may thus have both stronger problem-answer associations and stronger problem-procedure associations, resulting in higher retrieval and procedural efficiencies, respectively. Because practice enhances procedural efficiency, procedural strategies can be maintained as alternatives of equal value as retrieval strategies. Distributions of problem-answer associations and problem-procedure associations should thus be viewed as continuously dynamic, rather than reaching a final static state (LeFevre et al., 1996a).

But why were the effects of arithmetic practice significant for multiplication and not for addition? One explanation for this operation-dependent effect is based on the fact that strategy efficiencies differ between addition and multiplication. More specifically, compared to retrieval, procedures are less efficient for multiplication than they are for addition (Campbell & Xue, 2001). Because people always try to select the most efficient strategy (Siegler & Shipley, 1995), they will especially limit the use of procedures in order to solve multiplications. Otherwise stated, they will try to use retrieval more often, and more so for multiplication than for addition. Daily arithmetic practice may have enhanced this effect. Indeed, more-practiced students did retrieve multiplications more frequently than did less-practiced students. Moreover, since procedures are less efficient for multiplications than for additions, multiplication procedures are more susceptible to amelioration than addition procedures are. Two other factors may account for the operation-dependent effect on strategy efficiency as well: (a) the more frequent usage of multiplication in arithmetic and scientific classes and (b) the more declarative nature of multiplication (Roussel et al., 2002). Consequently, the more-practiced students may have built up stronger problem-answer and problem-procedure associations for multiplication than for addition, resulting in higher retrieval efficiency and higher procedural efficiency, respectively.

ARITHMETIC SKILL

Arithmetic skill was measured by means of a pen-and-paper test, and thus differed from the practice measure which was based on the number of arithmetic and scientific hours per week. According to LeFevre et al. (1996a), arithmetic skill can be viewed as a continuum from novice to expert, with high-skill participants retrieving arithmetic facts more frequently and more efficiently than low-skill participants (see also Ashcraft, Donley, Halas, & Vakali, 1992; Kaye, 1986; Kaye, de Winstanley, Chen, & Bonnefil, 1989; LeFevre & Kulak, 1994); these effects were observed in the present study as well. Indeed, regression analyses showed that arithmetic skill was the only individual trait that was highly predictive of strategy selection and strategy efficiency in both addition and multiplication. One may thus conclude that both strategy selection and strategy efficiency are potentially important indexes of individual differences in arithmetic skill.

Differences across our participants may also be compared with differences across cultures, as East-Asian and North-American students differ in retrieval efficiency, procedural efficiency, and retrieval use as well (LeFevre & Liu, 1997). LeFevre and Liu (1997) however, do not believe that these differences between East-Asians and North-Americans are due to overall differences in arithmetic skill. They state that if arithmetic skill differences were the cause, “a comparison between any groups of less- and more-skilled individuals would yield similar patterns of effects” (p. 51). LeFevre and Liu (1997) therefore supposed that fundamental differences in the organization of basic arithmetic facts in memory were responsible for the observed differences between East-Asian and North-American students. In the present study, we did observe comparable patterns across our single-culture participants as have been observed across cultures (i.e., differences in both strategy selection and strategy efficiency). This result might imply that our single-culture participants differed in the organization of basic arithmetic facts in memory, an issue that merits future research.

Finally, it should be noted that the influence of arithmetic skill was greater than the influence of daily arithmetic practice. One may thus argue that arithmetic skill largely determines simple-arithmetic performance, with some space left for another, more experience-based individual trait as daily practice. To what extent arithmetic skill may depend on variable factors such as daily arithmetic practice and on more 'stable' factors such as general intelligence, is an issue that future research may pursue. Future research may also try to disentangle effects of arithmetic skill and effects of arithmetic practice, possibly by controlling arithmetic skill and manipulating the amount of practice.

THE FREQUENCY OF CALCULATOR USE

Although cultural differences in the frequency of calculator use have been found (LeFevre & Liu, 1997), this study was the first to relate differences in calculator use to arithmetic performance within one single culture. A first observation was that the frequency of calculator use increased as children grow older (see also Campbell & Xue, 2001). One may further assume that this frequency will increase even more as nowadays most people (even children) always have cell phones at hand, which are also suited to calculate answers to various arithmetic problems. The frequent use of calculators may pose problems, however, as the present study showed a negative relationship between calculator use and strategy efficiency: Both retrieval and procedural strategies were executed less efficiently as the frequency of calculator use grew larger. One caution has to be made with this assertion, however: regression analyses show the relationship between calculator use, on the one hand, and strategy selection and strategy efficiency, on the other, which is insufficient to infer the direction of causality. Future studies are thus needed to investigate whether frequent calculator use results in poorer arithmetic performance or whether students poorer in mental arithmetic are more inclined to use the calculator.

GENDER DIFFERENCES

Although we did not really expect to observe differences between our male and female young adults' simple-arithmetic performance, regression data indicated that boys were somewhat more efficient in retrieving multiplication facts from memory than girls were. As any cognitive difference between males and females is of scientific and social importance, this issue is discussed further in the following. Royer and his colleagues (1999a) also observed more efficient retrieval use in boys than in girls, but the present study was the first one to observe comparable effects in young adults. It has been argued however, that any gender difference in retrieval efficiency is not likely to be directly based on biological factors such as sex hormones or primary memory systems (e.g., Geary, 1999; Royer et al., 1999a). Geary (1999) rather proposed that sexual selection might *indirectly* influence gender differences in mathematical cognition. More specifically, he proposes that the cognitive systems enabling movement in and the representation of three-dimensional space are more highly elaborated in males than in females. As these brain systems may support mathematical cognition, they may account for the observed differences in mental arithmetic between males and females. Further testing confirmed that the male advantage in mental arithmetic was mediated by gender differences, favoring males, in both spatial abilities and retrieval efficiency (Geary et al., 2000b).

It should be noted, however, that arithmetic performance is not only determined by retrieval efficiency, but also by procedural efficiency and retrieval frequency, and based on present results males and females do not differ in these latter aspects. Further research is needed to confirm or reject gender differences in speed of retrieval, and to further investigate sources of these and other differences between males and females in various cognitive domains.

OTHER VARIABLES

Although the present study examined various individual differences, others still remain unexamined. We chose to focus on cognitive differences across people, who may – of course – also vary regarding their emotionality. Previous studies indeed showed that strategic aspects of simple arithmetic might be influenced by people's affect towards mathematics (LeFevre et al., 1996a) and by their math anxiety (Ashcraft & Faust, 1994; Ashcraft & Kirk, 2001; Faust, Ashcraft, & Fleck, 1996). If we assume that high-anxiety students choose curricula with fewer hours of mathematics whereas low-anxiety students choose curricula with more hours of mathematics, such emotional factors should be seen as confounding variables, because they were not controlled in the present study. Besides such effects of self-selection, the effects of parental selection should not be denied either. It can be assumed that parents influence their children to choose a study curriculum with low or high amounts of arithmetic and scientific classes per week. It would be worthwhile to investigate whether arithmetic practice interacts with the enthusiasm with which students have chosen their study curriculum. Finally, general intelligence might have been an additional confounding variable. As with the factors mentioned earlier, it is difficult to control for intelligence. The present study thus cannot exclude effects of general intelligence completely. Future research should aim at disentangling effects of emotional, parental and cognitive factors (including general cognitive ability) on strategic performance in simple arithmetic.

IMPLICATIONS FOR SIMPLE-ARITHMETIC MODELS

When solving simple-arithmetic problems, multiple strategy use appears to be common in young adults (see also Campbell & Xue, 2001; Hecht, 1999; LeFevre & Liu, 1997; LeFevre et al., 1996a, 1996b). Because most present models of mathematical cognition simply account for retrieval-only data, such models cannot account for all simple-arithmetic performances in adults. Some form of a multiple-strategy model is required. The majority of current models cannot explain multiple strategy use, have not included experiential or educational factors in order to explain performance differences across participants, or both. Measures of associative strength for example, may be associated with individual differences such as daily arithmetic practice or arithmetic skill. However, as Hecht (1999) noted, most associative network models assume that the problem-answer associations are quite similar among adults because adults all share common experiences that influence the formation of their associative network. These ‘common experiences’ refer to how math facts are practiced and studied in childhood. The present study, however, showed that effects of experience do not stop after childhood; practice still influenced simple-arithmetic performance in young adults. Incorporating the on-going development of human being would be a true enrichment for models of mental arithmetic. Together with other researchers (e.g., LeFevre et al., 1996a; LeFevre & Liu, 1997) we thus believe that theories of mathematical cognition should include the possibility that reasonably skilled adults with different experiential backgrounds may vary in patterns of performance. Therefore, adequate models of mental arithmetic must make additional assumptions to account for both retrieval and nonretrieval responding, and must incorporate a role for individual differences and their consequences on arithmetic performance.

One model that – after some small modifications – would be able to account for our and others’ data (e.g., LeFevre et al., 1996a) is the adaptive strategy choice model (ASCM; Siegler & Shipley, 1995). According to this model, problems, strategies, and problem-strategy strengths are stored in a database. Each time a problem is solved, information about the strategy efficiency (e.g., time and accuracy) is added to the database, which can be modified accordingly. The probability of choosing a particular strategy is based on the strategy's strength relative to the strength of all other strategies. Furthermore, each time a particular strategy is used to solve a problem, the association between the problem and the strategy used gains in strength. In such a model, adults are no longer assumed to uniquely rely on retrieval. If we further assume that the amount of daily arithmetic practice influences problem-strategy associations (e.g., practice might strengthen the problem-retrieval association and weaken the problem-procedure association), ASCM would predict performance differences across participants with different amounts of practice. Therefore, we do not propose that the basic structure of the models should be changed, but that they should be updated by not only including problem features (e.g., problem size), but also strategy selection characteristics (e.g., single or multiple strategy use), and individual traits (e.g. arithmetic skill, arithmetic practice), in order to develop complete models of mental arithmetic that are able to explain various effects in simple-arithmetic performance. Previous studies already suggested some ideas for modifications to present models (e.g., Hecht, 1999; LeFevre et al., 1996a, 1996b).

CONCLUSIONS

Even though it has been suggested that experiential factors may influence the performance at early stages of learning whereas problem size would influence the performance of highly practiced adults (e.g., Geary, 1996), we believe that the influence of experiential factors will never disappear completely. Practice influences strategy selection and strategy efficiency in children (e.g., Siegler, 1986) as well as in young adults (present study). We therefore believe that future research should not confine itself to the investigation of skill variables but should also concentrate on other individual differences such as daily arithmetic practice. More general, and in agreement with other researchers (e.g., Kirk & Ashcraft, 2001; LeFevre et al., 1996b), we believe that investigations of individual differences and their relationship to online performance should be a priority for the field.

ADDENDUM

As noted in the introduction (Chapter 1), the original result section of this study reported in Chapter 2 did not exist merely of regression analyses. Standard analyses of variance (ANOVAs) and ex-Gaussian analyses were included as well. Although these analyses did not reach publication, we believe they provide an additional view on the data. They enabled us (a) to test whether daily arithmetic practice influenced the problem-size effect (i.e., a practice x size interaction), and (b) to directly compare addition and multiplication performance. As regression analyses were performed for each operation separately, such a comparison was not possible there. Since the results obtained by the standard analyses of variance were similar to those obtained by the ex-Gaussian analyses, we will only present the ex-Gaussian analyses. An additional merit of these analyses is that they provide very detailed information about strategy efficiency since they avoid the potential bias provoked by asking participants to describe their solution strategies (Penner-Wilger, Leth-Steensen, & LeFevre, 2002).

Note that the ex-Gaussian analyses were based on exactly the same data set as used for the regression analyses presented above[6]. However, problem size was now defined categorically and not continuously (as it was in the regression analyses). For both addition and multiplication problems, small problems were defined as those with products smaller than 25, whereas large problems were defined as those with products larger than 25 (cf. Campbell & Xue, 2001). Furthermore, five groups of participants were created: one with 3 hours of daily practice, one with 4 hours of daily practice, one with 9 or 11 hours of daily practice, one with 13 hours of daily practice, and one with 15 hours of daily practice.

[6] However, differences related to arithmetic skill, gender, and calculator use were not investigated now.

EX-GAUSSIAN ANALYSES

Performing ex-Gaussian analyses is a relatively novel method in which the shapes of RT distributions are examined (see also Penner-Wilger et al., 2002). More specifically, we used the ex-Gaussian distributional model to obtain quantitative measures of distributional shape (Heathcote, 1996). This distribution model provides a good fit to most RT data (Hockley, 1984; Hohle, 1965; Ratcliff & Murdock, 1976). As the ex-Gaussian distributional model consists of a normally distributed (Gaussian) leading edge and an exponentially distributed tail, its distributional shape can be summarized in three quantitative measures: *mu*, the mean of the normal component, *sigma*, the standard deviation of the normal component, and *tau*, the mean of the exponential component. Since these measures can be used to describe the full profile of a set of RTs, they provide much more information than can be obtained from mean RT measures alone (Penner-Wilger et al., 2002). The *mu* value stands for the leading edge of the RT distribution and is composed of the relatively faster RTs, and thus largely reflects retrieval efficiency. The *tau* value, in contrast, stands for the tail, which is composed of the relatively slower RTs. As nonretrieval strategies generally produce slower responses, procedural efficiency should be reflected in this *tau* value, although the *tau* value might also reflect slow retrieval processes.

The three parameters of the ex-Gaussian distribution were estimated by means of maximum likelihood. A robust optimizer (an implementation of the simplex algorithm, Nelder & Mead, 1965, available in the statistical package R) was used to maximize the (log) likelihood function. Trials with RTs larger than 10 seconds were removed from analyses. For each fit, the procedure was repeated 10 times with random starting points in order to avoid local maxima solutions. This was done for 240 datasets (i.e., 2 operations x 2 sizes x 60 participants); the number of data points per fit varied from 42 to 58 (mean = 54.2; standard deviation = 2.3). For each fit, a goodness-of-fit measure for the ex-Gaussian distribution was computed in terms of a chi-square statistic. This statistic showed that only 7% of the tests indicated a poor approximation of the RT distribution by the ex-Gaussian model. Nevertheless, ex-Gaussian values with poor fits to the distributional data were included in the analyses, because these values still provide meaningful quantitative information about the general shapes of the distributions (see also Penner-Wilger et al., 2002).

Analyses on mu and sigma. A 5 (group) x 2 (size) x 2 (operation) ANOVA was performed on the ex-Gaussian parameter *mu* (see Table 4). The mean of this ex-Gaussian normal component was larger for large than for small problems, $F(1,55) = 101.8$, $MSe = 17529$ (823 vs. 646 msec), indicating a problem-size effect in retrieval RTs. Further, *mu* was also larger for multiplication than for addition, $F(1,55) = 2.8$, $MSe = 11344$, $p = .099$ (746 vs. 722 msec), whereas the interaction between size and operation just failed to reach significance, $F(1,55) = 2.6$, $MSe = 8183$, $p = .11$, with larger problem-size effects for multiplication than for addition (197 vs. 58 msec). Although *mu* did not differ significantly across groups, $F(4,55) = 1.2$, $MSe = 74094$, a further planned comparison showed significantly larger *mu* values for the three less-practiced groups as compared to the two more-practiced groups, $t(55) = 1.6$, indicating more efficient retrieval in the latter group than in the former one. This effect tended to be larger for large problems than for small problems, $t(55) = 1.5$, $p = .07$. Thus, although the problem-size effect in *mu* was significant in all groups (all p's $< .001$), it decreased as the practice level increased. The variability of the ex-Gaussian normal component (*sigma*, see Table 5) showed only one significant effect, namely that of size, with *sigma* being larger for large problems than for small problems, $F(1,557) = 33.7$, $MSe = 5652$ (121 vs. 63 msec).

Table 4
Means of the ex-Gaussian *mu* value as a function of group, size, and operation. Standard errors are shown between brackets

	Addition		Multiplication	
Daily arithmetic practice	Small	Large	Small	Large
3	626 (24)	789 (44)	634 (25)	829 (65)
4	668 (23)	858 (43)	680 (24)	915 (63)
9-11	667 (29)	846 (54)	670 (30)	929 (80)
13	619 (29)	747 (54)	618 (30)	749 (80)
15	638 (31)	768 (57)	636 (32)	799 (84)

Analyses on tau. The mean of the ex-Gaussian exponential component (*tau*, see Table 6) was significantly larger for multiplication than for addition, $F(1,55) = 53.3$, $MSe = 95565$ (522 vs. 249 msec) and larger for large problems than for small problems, $F(1,55) = 66.6$, $MSe = 66125$ (525 vs. 246 msec), indicating a significant procedural problem-size effect. Size and operation interacted with each other, $F(1,55) = 23.6$, $MSe = 42706$, with larger procedural problem-size effects in multiplication than in addition (412 vs. 145 msec). Most importantly, however, was the main effect of group, $F(4,55) = 3.5$, $MSe = 196334$, showing significantly larger *tau* values as the practice level decreased. Moreover, this effect interacted with problem size, $F(4,55) = 2.9$, showing larger procedural problem-size effects as the practice level decreased. Hence, all participants needed more time when executing procedures on large problems than on small problems, although this effect became smaller with practice.

Table 5
Means of the ex-Gaussian *sigma* value as a function of group, size, and operation. Standard errors are shown between brackets.

	Addition		Multiplication	
Daily arithmetic practice	Small	Large	Small	Large
3	49 (9)	98 (16)	30 (8)	132 (33)
4	67 (8)	100 (16)	49 (8)	150 (32)
9-11	79 (11)	125 (20)	60 (10)	145 (41)
13	76 (11)	111 (20)	79 (10)	108 (41)
15	70 (11)	124 (21)	69 (11)	114 (43)

Vincentized group RT distributions. To visualize the results described above, the Vincent averaging technique recommended by Ratcliff (1979; see also Heathcote, 1996; Penner-Wilger et al., 2002) was used to derive RT distributions for all groups and for both small and large problems. As can been seen in Figures 1a – 1c, the RT distributional patterns differed across the student groups. As the amount of daily practice increased, the distributions became more peaked, with smaller positive skews. These differences were more apparent for large problems than for small problems. As problem size increases, all groups show a shift in the leading edge and an increase in the size of the tail. Though, these differences were larger for the less-practiced groups than for the more-practiced groups.

Table 6
Means of the ex-Gaussian *tau* value as a function of group, size, and operation. Standard errors are shown between brackets.

	Addition		Multiplication	
Daily arithmetic practice	Small	Large	Small	Large
3	212 (22)	371 (47)	346 (48)	946 (147)
4	187 (21)	447 (45)	401 (46)	977 (143)
9-11	218 (26)	365 (57)	384 (59)	839 (180)
13	113 (26)	214 (57)	220 (59)	434 (180)
15	150 (28)	211 (60)	232 (62)	443 (190)

Figure 1a
Probability density histograms for RT distributions for students with 3 practice hours.

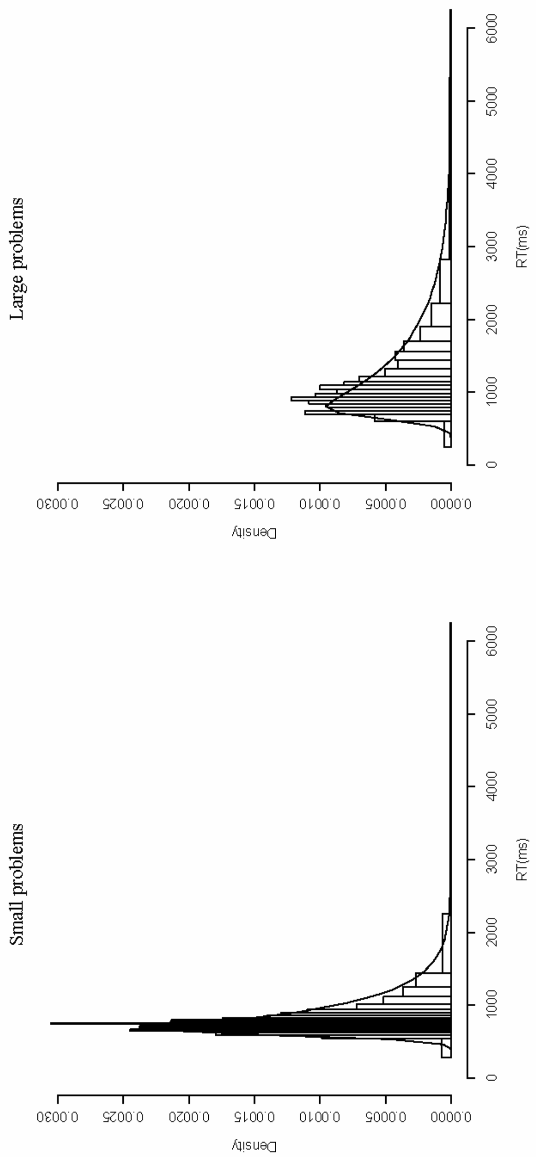


Figure 1b
Probability density histograms for RT distributions for students with 5 practice hours.

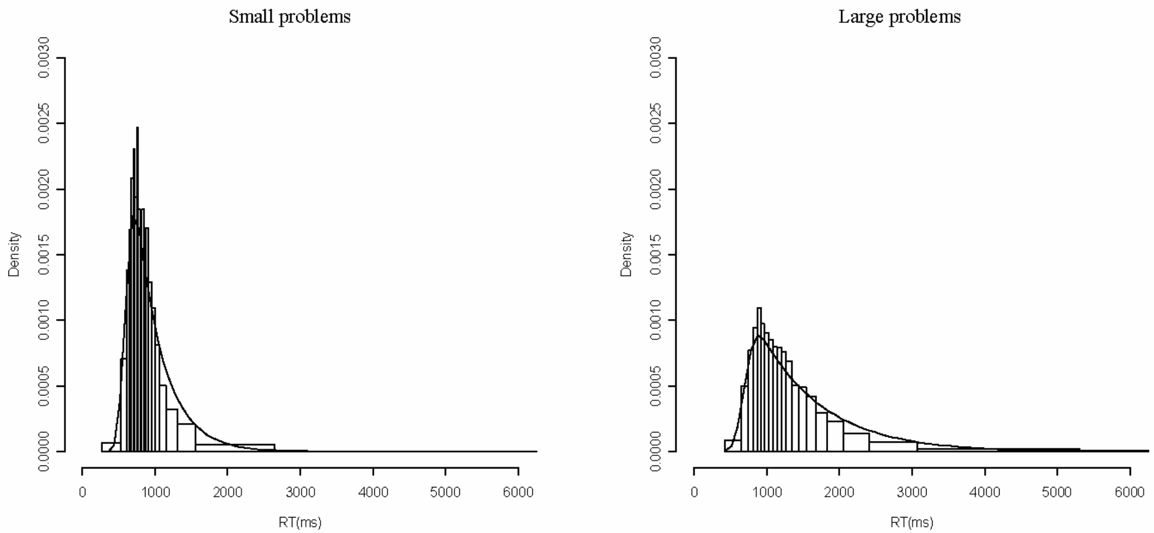
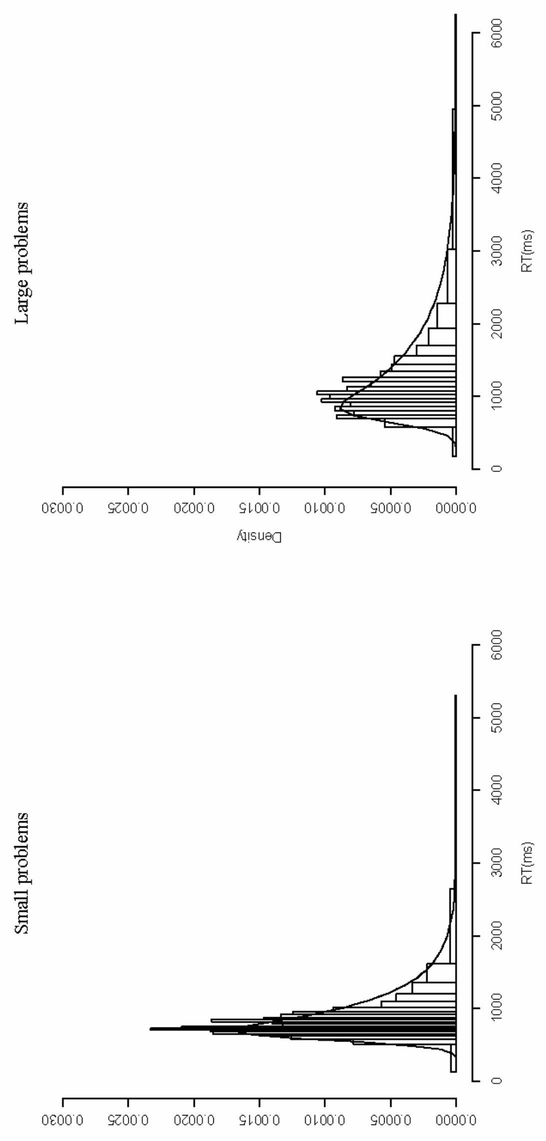


Figure 1c
Probability density histograms for RT distributions for students with 9-11 practice hours.



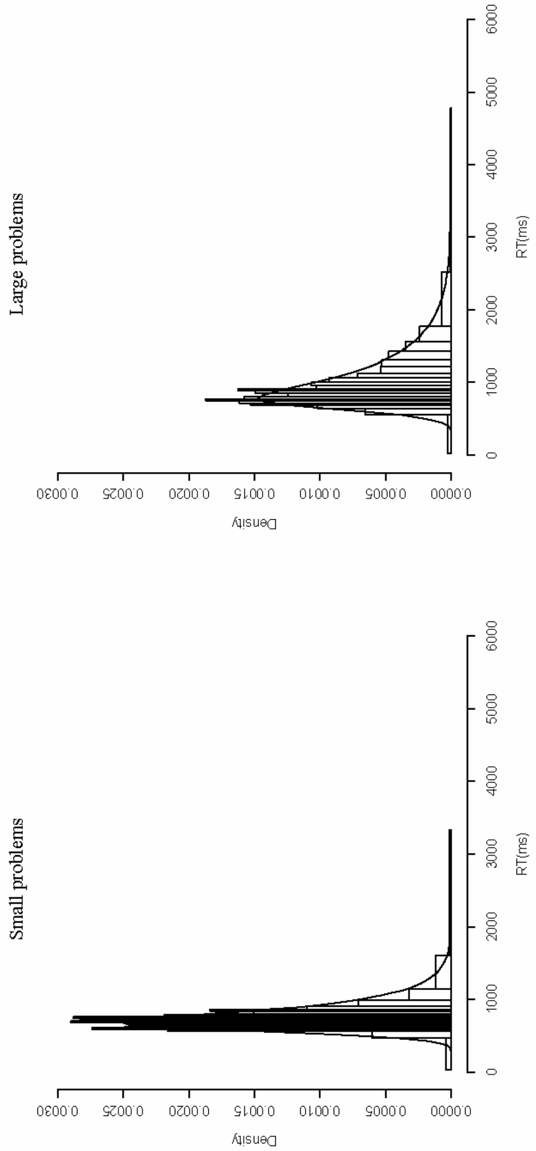
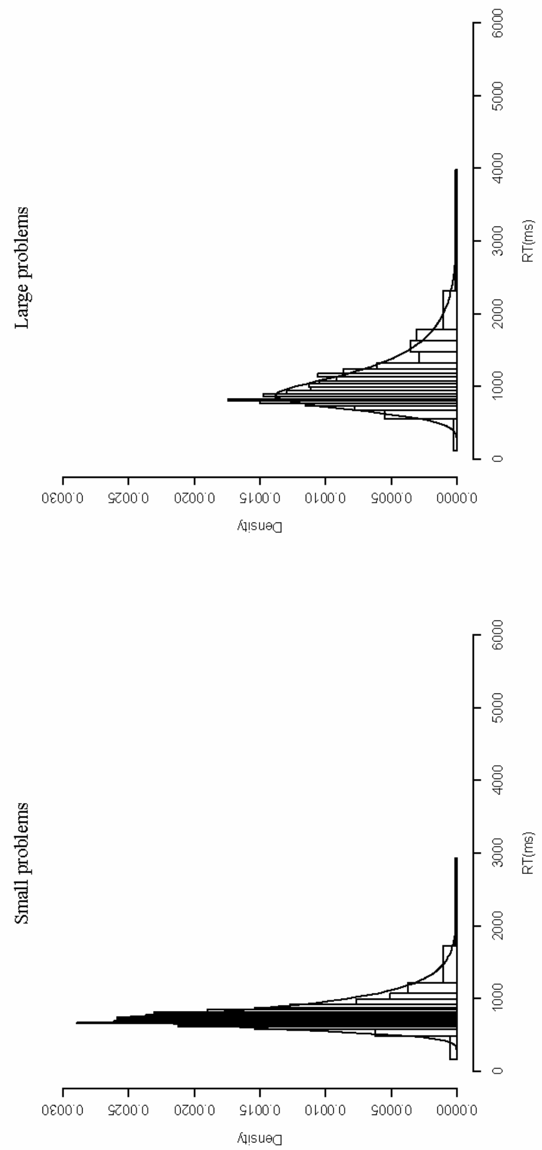


Figure 1d
Probability density histograms for RT distributions for students with 13 practice hours.

Figure 1e
Probability density histograms for RT distributions for students with 15 practice hours.



DISCUSSION

It is important to note that in the ex-Gaussian analyses, trials were not separated on the basis of the participants' strategy reports. Instead, all trials were analyzed collectively, thereby avoiding possible biasing effects caused by strategy reports. Though, using such a different technique, the ex-Gaussian analyses confirmed the practice effects observed in the regression analyses. The values of both *mu* and *tau* differed across groups. More-practiced participants reached higher retrieval efficiencies and higher procedural efficiencies than did less-practiced participants.

More importantly, the ex-Gaussian analyses also provided information that was *not* obtained in the regression analyses. First, operation influenced retrieval efficiency and procedural efficiency; both strategy efficiencies were higher in addition than in multiplication. These results are in agreement with previous studies (e.g. Campbell & Xue, 2001).

Second, daily arithmetic practice influenced the problem-size effect. Both the retrieval problem-size effect and the procedural problem-size effect were significant for all participants; *mu* and *tau* values were larger for large problems than for small problems. However, the differences between small and large problems decreased as daily arithmetic practice increased. The retrieval problem-size effect and the procedural problem-size effect thus *de*creased as practice level *in*creased. The patterns observed in the vincentized RT distributions corresponded to these effects in *mu* and *tau*, since they showed that both the leading edge and the size of the tail differed across groups and across problem sizes.

To conclude, more efficient retrieval use and more efficient procedural use on small problems than on large problems not only contributed to the problem-size effect; these differences between small and large problems also contributed to the differential magnitudes of the problem-size effect across more- and less-practiced participants. This latter result could not be obtained by the regression analyses.

It is interesting to compare our results with those obtained in other cultures. As noted above, Penner-Wilger et al. (2002) investigated whether the locus of the problem-size effect differed across cultures. Using the data obtained by LeFevre and Liu (1997), but applying the ex-Gaussian distributional model, Penner-Wilger and colleagues showed that the problem-size effect in East-Asians was purely a memory-retrieval effect (cf. a problem-size effect in *mu*), whereas in North-Americans, it was an effect of both retrieval efficiency and procedural efficiency (cf. a problem-size effect in both *mu* and *tau*). The participants in the current study showed a problem-size effect in both *mu* and *tau*; their pattern thus resembled North-Americans but did not resemble East-Asians. Hence, the problem-size effect in both North-American and West-European students can be attributed to less efficient retrieval use and less efficient procedural use on large versus small problems, whereas the problem-size effect in East-Asian students can be fully attributed to less efficient retrieval on large versus small problems.

CHAPTER 3
PRACTICE EFFECTS ON STRATEGY SELECTION AND STRATEGY EFFICIENCY IN SIMPLE MENTAL ARITHMETIC

***Manuscript under revision*[1]**

Two experiments were conducted to investigate the effects of practice on strategy selection and strategy efficiency in mental arithmetic. Participants had to solve simple addition or multiplication problems, after having received 0, 3, or 6 practice sessions (Experiment 1), and before and after having received 3 practice sessions (Experiment 2). Although practice consisted of simple-arithmetic problems only, a test measuring complex-arithmetic performance was administered after the practice sessions as well. Results showed practice effects on retrieval frequency, retrieval efficiency, and procedural efficiency. Long-term practice effects appeared to be both strategy-specific (i.e., only for procedural strategies) and operation-specific (i.e., only for multiplication problems), and they generalized over problem complexity. Implications of the present results for mathematic cognition and its modeling are discussed.

---

[1] This paper was co-authored by André Vandierendonck.

INTRODUCTION

Daily, we use several numeric competencies, such as subitizing small quantities, estimating large quantities, calculating new quantities, et cetera. Some of these basic competencies, such as subitizing and estimating, seem to be innate to human infants (e.g., Butterworth, 1999; Dehaene, 1997; Spelke & Dehaene, 1999). The mastery of more advanced numerical skills such as calculation, in contrast, must be acquired through education, learning, and practice[2]. Between the ages of 2 and 4, children learn to count verbally (e.g., Gelman & Gallistel, 1978; Wynn, 1990). Once they master counting, children generalize the counting procedure to larger numbers, apparently without upper bound and without explicit training. During their elementary school years, children learn a set of basic arithmetic facts and calculation procedures. Slow, deliberate and effortful procedures (such as step-by-step counting) are replaced by fast, efficient, and less effortful calculation processes (such as memory retrieval; e.g., Ashcraft, 1982; Fuson, 1982, 1988; Siegler, 1988b; Steel & Funnell, 2001). When reaching adult age, until recently, people were supposed always to use memory retrieval to solve simple-arithmetic problems such as 3 + 5 and 6 x 7 (e.g., Ashcraft, 1992; McCloskey, 1992). However, more recent studies showed that this is not absolutely true: even skilled adults are not always able to retrieve simple-arithmetic facts from their memory (e.g., LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). Indeed, many adults use nonretrieval (procedural) strategies such as counting (e.g., 6 + 3 = 6 + 1 + 1 + 1) and transformation (e.g., 6 + 7 = 6 + 4 + 3) to solve simple-arithmetic problems. Other quite surprising results have been found concerning the arithmetic abilities of normal educated adults. Geary and colleagues (1996b,

---

[2] In the present paper, the term ‘training’ is seen as the explicit training of a particular strategy, whereas the term ‘practice’ is seen as exercising through repetition, without enforcing a particular strategy.

1997) showed declines in mental arithmetic performance across successive North-American generations. Comparable results were obtained by Mulhern and Wylie (2004), who showed that performance levels of psychology students on core mathematical skills (such as calculation) dropped devastatingly between 1992 and 2002.

Notwithstanding the importance of mental arithmetic in daily life, and the decline of mathematic skill the last few years, few studies so far investigated the effects of practice on strategy selection and strategy efficiency. Practice effects in simple arithmetic have been studied in children (see Kroesbergen & Van Luit, 2003, for a review), and in brain-damaged adults (e.g., Whetstone, 1998), but nearly not in healthy adults (but see Fendrich, Healy, & Bourne, 1993; Pauli, Bourne, & Birbaumer, 1998; Rickard, Healy, & Bourne, 1994). Four questions may be raised: (1) Does practice increase the use of direct memory retrieval (i.e., a change in strategy selection)?, (2) Does practice increase the speed with which retrieval and procedural strategies are executed (i.e., an increase in strategy efficiency)?, (3) Does practice reduce the performance differences between small and large problems (i.e., the problem-size effect)?, and (4) Do practice effects transfer to other operations, other sizes, or other arithmetic problems?

Concerning the first question (i.e., whether practice influences strategy selection), the most prevailing assumption is that practice will inevitably lead to an augmented usage of retrieval. This assumption is based on the distribution of associations model (Siegler & Shrager, 1984) and the instance theory of automatization (Logan, 1988). The distribution of associations model (Siegler & Shrager, 1984) states that the encoding of a problem results in the activation of a set of response candidates. The activation of each candidate depends on the acquired problem-answer strength. It is further assumed that there is a direct relation between the activation level and the probability of retrieval. Answers with a high associative strength will be retrieved, but if the problem-answer associative strength does not exceed a predefined confidence criterion, a procedural strategy will be used in order to solve the problem. As continued practice strengthens problem-answer associations, retrieval will be used more frequently, resulting in a concomitant decrease in procedural strategy use. In a later version of this model, the adaptive strategy choice model (ASCM, Siegler & Shipley, 1995), selection of an arithmetic strategy depends on its relative efficiency (i.e., speed and accuracy). As a result of practice, problem-strategy associations increase, and this increase is as large for both retrieval and procedural strategies. This model thus not only predicts an increase in the use of retrieval strategies, but also an increase in the use of more efficient procedural strategies relative to less efficient ones. However, in the end, extensive practice should result in exclusive retrieval use.

According to Logan's (1988) instance theory of automatization, each encounter with a stimulus initiates a race between procedural and retrieval strategies. In the beginning, the race is predominantly won by procedures. However, as each problem encounter is encoded and stored in long-term memory, practice enhances retrieval speed but not procedural speed. Consequently, as practice progresses, the retrieval strategy will win the race. Otherwise stated, practice enhances the amount of automatization, which reflects a transition from performance based on procedural strategies to performance based on memory retrieval.

However, since the use of procedural strategies persists even in skilled adults, LeFevre et al. (1996a) maintain that practice will *not always* lead to increased usage of retrieval. According to these authors, practice can also lead to the automatic activation of procedural strategies. In this view, associations between a specific problem and a procedure are created and strengthened by the successful use of such procedural strategies. Therefore, when people encounter that problem, they will automatically activate a procedure to solve that problem, without (or before) trying to retrieve it from long-term memory. Practice will then not solely lead to the replacement of procedures by fact retrieval, but also to the replacement of less efficient procedures by more efficient procedures. This view (see also Baroody, 1983, 1984, 1985) implies that the availability of efficient procedural skills would avoid the necessity of memorizing all the basic number combinations. Consequently, many simple-arithmetic problems might continue to be solved by using procedural strategies.

Up until now, evidence concerning practice effects on strategy selection has only been shown for the alphabet arithmetic task (e.g., Brigman & Cherry, 2002; Compton & Logan, 1991; Hoyer, Cerella, & Onyper, 2003; Logan, 1988; Logan & Klapp, 1991, Zbrodoff, 1999), for pseudo-arithmetic tasks (e.g., Onyper, Hoyer, & Cerella, 2006; Rickard, 1997; Touron, Hoyer, & Cerella, 2004), and for arithmetic word problems (e.g., Lewis, 1989). The current study aims at investigating practice effects on strategy selection in simple-arithmetic tasks.

The second question concerns practice effects on strategy efficiency. According to the distribution of associations model (Siegler & Shrager, 1984), successive correct practice trials strengthen the link between a problem and its answer. Since the time to retrieve and produce an answer is proportional to the activation level of the corresponding answer node, this theory predicts that practice will increase retrieval efficiency. The later version of this model, the ASCM (Siegler & Shipley, 1995), predicts an increase in the efficiency with which *each* strategy is executed, and thus predicts higher efficiencies in both retrieval and procedural strategies. Similar predictions can be made based on Rickard's (1997) theory of skill acquisition, the component power law (CMPL). This theory involves separate strategy nodes for computation and retrieval that both can be strengthened.

The instance theory of automatization (Logan, 1988), in contrast, assumes that the finishing times for procedures stay the same while the finishing times for retrieval decrease. This theory thus predicts practice effects on retrieval efficiency but not on procedural efficiency. Finally, Baroody (1983, 1984, 1985) predicts rather the opposite. According to his procedure-based theory, skill acquisition is based on the replacement of slow procedural strategies by faster, more automatic procedural strategies rather than the replacement of procedures by direct memory retrieval. He thus predicts stronger practice effects on procedural efficiency than on retrieval efficiency.

Effects of practice on strategy efficiency have been observed in standard arithmetic problems (e.g., Campbell, 1987a; Pauli et al., 1998), in pseudo-arithmetic tasks (e.g., Onyper et al., 2006; Touron et al., 2004), and in alphabet arithmetic tasks (e.g., Hoyer et al., 2003; Klapp, Boches, Trabert, & Logan, 1991). However, since many studies did not include strategy reports, it is not clear whether the faster response times were due to changes in strategy efficiency (i.e., faster strategy execution) or to changes in strategy selection (i.e., more frequent use of faster strategies)[3]. Moreover, even when authors maintain that practice enhanced strategy efficiency, they could not point out *which* strategies benefited most from practice. The present study aims to investigate practice effects on retrieval efficiency and procedural efficiency separately.

Thirdly, the present study aims to test why the problem-size effect is modified by practice. The problem-size effect refers to the observation that large problems such as 8 x 9 take longer to solve than small problems such as 2 x 3. It has been shown that the problem-size effect decreases as a result of practice (e.g., Fendrich et al., 1993; Pauli et al., 1998). The problem-size effect may decrease in three ways: more frequent retrieval use for large problems, more efficient retrieval use for large problems, and more efficient

[3] Some studies (e.g., Compton & Logan, 1991; Rickard, 1997) included strategy reports on *subsets* of trials, e.g., on one sixth of the trials. Logan and Klapp (1991) asked participants – at the end of the experiment – to estimate the percentage of trials on which they had used retrieval vs. counting strategies. Strategy reports on *all* trials have been used in alphabet arithmetic tasks (e.g., Hoyer et al., 2003) and in pseudo arithmetic tasks (e.g., Onyper et al., 2006; Touron et al., 2004).

procedural use for large numbers (Campbell & Xue, 2001). The present study tests whether practice influences all three of them.

The final question raised concerns the transfer of practice to other operations, other sizes, or other arithmetic problems. According to associative network theories (e.g., Campbell, 1987a; Campbell & Graham, 1985), the instance theory of automatization (Logan, 1988), and the identical elements model (Rickard et al., 1994), practice effects are item-based rather than process-based. This implies that practice involves learning specific responses to specific stimuli. Consequently, transfer to novel stimuli and situations should be inexistent.

According to the ASCM (Siegler & Shipley, 1995) and the CMPL theory (Rickard, 1997), in contrast, practice enhances both retrieval and procedural efficiencies. If it is further assumed that procedures can be applied to several stimuli, these models predict that practice effects on simple-arithmetic problems will transfer to complex-arithmetic problems. This reasoning is also adopted in the procedure-based view of Baroody (1983, 1984, 1985), who entails that procedural strategy use is cognitively more economical than retrieval use because it can be used on multiple problems.

Previous studies on mental arithmetic reported transfer for highly related problems, such as commuted problems, but not for other problems or other operations (e.g., Campbell, 1987a; Fendrich, et al., 1993; Pauli et al., 1994; Rickard et al., 1994). Pauli et al. (1998) even did observe no overall transfer from practiced to new multiplication problems. Practice was thus item-specific and did not facilitate arithmetic performance on problems that were not practiced. More recently, Delazer et al. (2005) showed that transfer from old to new complex addition problems only occurred when procedural strategies had been practiced but not when direct memory retrieval had been practiced. We wondered whether the same strategy-dependent effect of transfer would be true when simple arithmetic is practiced.

The present study consists of two experiments, which were conducted in order to formulate an answer to the four questions outlined above.

EXPERIMENT 1

In this experiment, participants had to solve simple-arithmetic addition or multiplication problems. Both latency data and strategy reports were collected in order to investigate practice effects on strategy selection and strategy efficiency. On the assumption that small problems are usually solved very efficiently (cf. the problem-size effect), only large problems were practiced. Participants were given 0, 3, or 6 practice sessions of the subset of largest problems. After these practice sessions, a test session was administered, in which both small and large problems had to be solved. All participants also had to complete a test of complex arithmetic (i.e., the French kit) after the test and practice sessions.

METHOD

Participants. Sixty first-year psychology students (9 men and 51 women) at Ghent University participated for course requirements and credits. Their mean age was 19.0 years. They were randomly assigned to the cells of a 2 (Operation) x 3 (Practice) design.

Procedure. Each participant was tested individually in a quiet room for approximately 30, 45, or 60 minutes (dependent on the practice condition). Two tasks were given to each participant. The first one was the simple-arithmetic task, which consisted of simple additions (for one group of 30 participants) or simple multiplications (for another group of 30 participants). Within each group, 10 participants did not practice, 10 participants completed 3 practice sessions, and 10 participants completed 6 practice sessions. After the practice sessions, all participants also completed a test session. The second task was the French kit (French, Ekstrom, & Price, 1963), which consists of complex-arithmetic problems that have to be solved as quickly and accurately as possible. In the following, both tasks are described more extensively.

Stimuli of the simple-arithmetic task consisted of simple addition and simple multiplication problems. Both addition and multiplication problems were composed of pairs of numbers between 2 and 9, with tie problems (e.g., 3 + 3) excluded. Problems involving 0 or 1 as an operand or answer were also excluded. This resulted in 56 addition problems (ranging from 2 + 3 to 8 + 9) and 56 multiplication problems (ranging from 2 x 3 to 8 x 9). Although all problems were presented in the test session, only the most difficult problems were presented in the practice sessions. The practice problems consisted of the 12 largest addition problems and the 12 largest multiplication problems. For addition, this selection included all problems with a sum ranging from 14 to 17. For multiplication, this selection included all problems with a product ranging from 45 to 72. Definition of small and large problems was also based on this selection: small problems were defined as the not-selected problems (i.e., the 44 smallest ones), whereas large problems were defined as the selected problems (i.e., the 12 largest ones). As noted before, there were three practice conditions: 0, 3, or 6 practice sessions. Within each practice session, all practice problems (i.e., the large ones) were presented twice, and in the test session all problems (small and large ones) were presented twice. All problems were presented in Arabic format and in a randomized order within one session.

A trial started with a fixation point, which appeared for 500 msec. Then the arithmetic task appeared in the center of the screen. The addition and multiplication problems were presented horizontally as dark-blue characters on a light-grey background, with the operation sign at the fixation point. The problem remained on screen until the participant responded. In order to avoid biasing conditions, no time deadline was set, since it has been shown that a fast deadline increases reported use of retrieval, especially for large problems (Campbell & Austin, 2002). A sound-activated relay was activated when participants spoke their answer aloud in a microphone, which was connected to a software clock (accurate to 1 millisecond). The use of a voice-key minimized general speeding effects in motor responses during practice. In previous research (e.g., Rickard et al., 1994) participants often had to type in the answer on the numeric key pad, so that improvements in motor aspects during practice might have influenced overall performance. Pauli et al. (1998) indeed showed decreases in both mental calculation time and motor response time across practice sessions. All invalid trials (e.g., failures of the voice-activated relay) were discarded, and (in the test session only) they returned at the end of the session.

On each trial, accuracy was registered online by the experimenter and feedback was presented to the participants, a green 'Correct' when their answer was correct, and a red 'Incorrect' when it was not. Participants were also told to report the strategy they used for each single problem. The reported strategy was recorded online by the experimenter by pressing a predefined number key on the keyboard. Participants could choose one of the four strategies described below (see e.g., Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler, Kirk, & Ashcraft, 2003): *(1) Remember: You solve the problem by just remembering or knowing the answer directly from memory; (2) Counting: You solve the problem by counting a certain number of times to get the answer; (3) Transformation: You solve the problem by referring to related operations or by deriving the answer from some known facts; and (4) Other: You solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem.* These four strategies were extensively explained by the experimenter, with examples of both addition or multiplication problems solved by each strategy as appropriate. It was emphasized that the presented strategies were not meant to encourage use of a particular strategy.

After the simple-arithmetic task, participants completed two arithmetic subtests of the French kit (French et al., 1963), one page of complex addition problems (e.g., 39 + 90 + 82) and one page of complex subtraction and multiplication problems (e.g., 48 x 7). Each page contained six rows of ten vertically oriented problems. Participants were given two minutes per page to solve the problems as quickly and accurately as possible. Scores were defined as the total number of correctly solved problems per test.

RESULTS

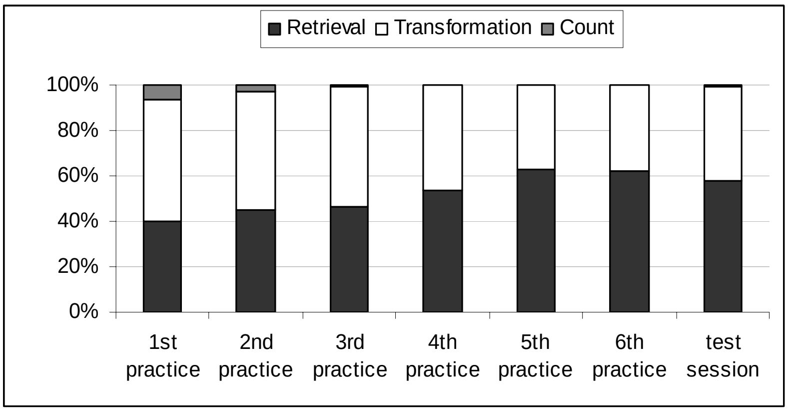
Across operations, 826 trials (6.8%) were spoiled due to failures of the sound-activated relay. Since all the invalid trials met in the test session returned at the end of this session, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 58 (0.5%). Further, all incorrect trials (367 trials) and all trials on which participants selected the 'Other' category (9 trials) were deleted. All response times (RTs) more than 4 standard deviations from each participant's mean (per operation) were discarded as outliers (88 trials). Finally, one participant (in the multiplication experiment with six practice sessions) was discarded due to voice key problems.

The results section is divided in two main parts. First, effects *during* the practice sessions are reported. Second, effects *after* the practice sessions (i.e., in the test session) are considered. Every analysis of variance (ANOVA) that has been executed was based on the multivariate linear model. All reported results are considered to be significant if $p < .05$, unless mentioned otherwise. Although no pre-practice test had been administered, we may assume that all observed effects were due to the manipulated variables, given that first-session RTs did not differ across groups (0x, 3x, 6x; $F < 1$).

Effects during the practice sessions. Only the participants who completed 3 or 6 practice sessions are included in the following analyses. Performance was grouped in four blocks for both practice levels. The first block comprises the performance in the first practice session for the 3x practice level, and the mean performance of the first and second practice session for the 6x practice level. The second block comprises the performance in the second practice session for the 3x practice level, and the mean performance of the third and fourth practice session for the 6x practice level. The third block comprises the third and final practice session for the 3x practice level, and the mean of both final practice sessions for the 6x practice level. The final block comprises the performance in the test session for both practice levels (large problems only). Although all analyses were performed on the 'block' variable, the tables show data of each single session as well. Since only the 12 most difficult problems were presented in the practice sessions, problem size was not included in the following analyses.

In order to obtain a first impression concerning practice effects on *strategy selection*, percentages of usage of all three strategies (on large problems only) across the six practice sessions are displayed in Figure 1. Although retrieval was the most frequently used strategy, procedural strategies (counting and transformation) were used too. The frequency of retrieval increased as practice proceeded, whereas frequencies of procedural strategy use decreased. For addition, the counting strategy was no longer used from the fourth practice session on. For multiplication, the counting strategy was no longer used from the fifth practice session on. In the last practice sessions, a 'steady state' was reached, with very frequent use of retrieval, infrequent use of the transformation strategy, and no use of the counting strategy. The figure also shows that the practice effects decreased in the test session (i.e., procedural strategy use increased and retrieval use decreased), where the very same (large) problems were presented among the other (small) problems.

Addition



Multiplication

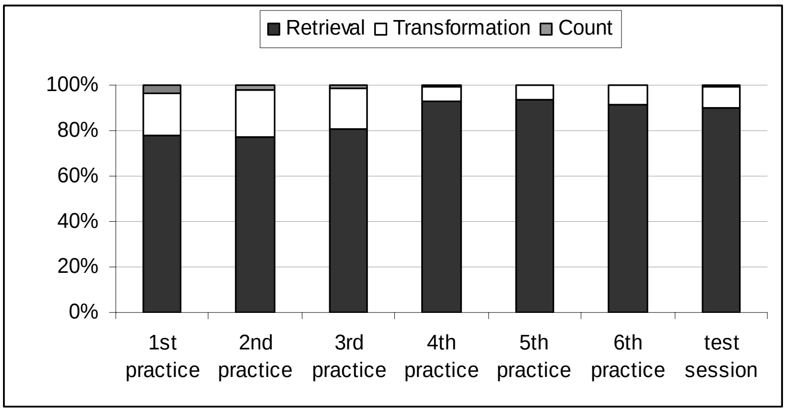


Figure 1
Percentage of use of retrieval, transformation, and counting strategies across sessions.
(Data of the 6x practice level only)

An ANOVA was run on percentages retrieval use, with block (first, second, third, test) as within-subjects variable, and practice level (3x vs. 6x) and operation (addition vs. multiplication) as between-subjects variable (see Table 1). The effect of block was significant, $F(3,33) = 5.67$. This effect consisted of both a linear effect, $F(1,35) = 6.46$, and a quadratic effect,

$F(1,35) = 9.98$. This indicated that reported use of direct retrieval increased linearly across practice blocks, but decreased again in the test session. Importantly, both the linear and the quadratic effect were significant for addition, $F(1,35) = 6.86$ and $F(1,35) = 12.44$, respectively, but not for multiplication (both $Fs < 1$). No main or interaction effects with practice level were observed.

Table 1 (continued on next page)
Percentages of retrieval use as a function of block[a], practice level, and operation. Standard errors are shown between brackets.

Addition	3x practice	6x practice	
	Across blocks (=sessions)	Across blocks	Across sessions
1st practice block	32 (9)	42 (9)	40 (7) 44 (9)
2nd practice block	51 (10)	50 (10)	48 (10) 53 (9)
3rd practice block	50 (9)	63 (9)	63 (9) 62 (9)
test session	39 (9)	58 (9)	58 (9)

[a] For the 3x practice group, each practice block comprised one practice session. For the 6x practice group, each practice block comprised two practice sessions. For completeness, percentages retrieval use are provided for each single practice session in the 6x practiced participants as well.

Table 1
Percentages of retrieval use as a function of block[a], practice level, and operation. Standard errors are shown between brackets.

Multiplication	3x practice	6x practice	
	Across blocks (=sessions)	Across blocks	Across sessions
1st practice block	73 (9)	77 (9)	78 (8) 77 (9)
2nd practice block	72 (10)	87 (10)	81 (11) 93 (9)
3rd practice block	71 (10)	93 (10)	93 (9) 91 (9)
test session	69 (9)	90 (10)	90 (10)

[a] For the 3x practice group, each practice block comprised one practice session. For the 6x practice group, each practice block comprised two practice sessions. For completeness, percentages retrieval use are provided for each single practice session in the 6x practiced participants as well.

Additional analyses were performed to analyze whether the increase in retrieval use was related to the initial amount of retrieval use. A median split on the percentage of retrieval use in the first practice session divided the participants in two groups. Nineteen participants (5 in the addition group and 14 in the multiplication group) were labeled as frequent-retrieval users (85% retrieval use in the first block) and twenty participants (15 in the addition group and 5 in the multiplication group) were labeled as infrequent-retrieval users (28% retrieval use in the first block). It was expected that the increase in retrieval use would be especially apparent in the latter group. This was confirmed by the data: the increase in retrieval use between the first practice block and the last practice block was significantly higher for infrequent-retrieval users (22%) than for frequent-retrieval users (3%), $t(37) = 3.38$. We also tested whether the decrease in retrieval use between the last practice block and the test session differed across participants. Surprisingly, this decrease was equally large (i.e., 5%) for both groups, $t < 1$. Thus, between the first practice session and the test session, the amount of retrieval use insignificantly *decreased* for frequent-retrieval users (-2%; $t < 1$), whereas it significantly *increased* for infrequent-retrieval users (+17%; $t(19) = 3.4$). By way of conclusion, these additional analyses showed that the increases in retrieval use were closely linked to the initial amount of retrieval use. More specifically, practice was most important for those participants who didn't use retrieval frequently. As retrieval is less frequently used in addition than in multiplication, practice influenced strategy selection more strongly in addition than in multiplication.

In order to test practice effects on *strategy efficiency*, an ANOVA was performed on RTs, with practice level (3x vs. 6x) and operation (addition vs. multiplication) as between-subjects variables, and block (first, second, third, test) and strategy (retrieval vs. procedural) as within-subjects variables[4] (see Tables 2a and 2b). Although accuracy is a component of strategy efficiency as well, accuracies were not analyzed given the very low

[4] Since (a) not all strategies were used across all the practice sessions, and (b) only RTs of the correctly solved problems were analyzed, for some subjects empty cells occurred in the practice level x operation x session/block x strategy ANOVA. We replaced these empty cells for each participant with the correct RT of the corresponding cell [i.e., the mean RT (over participants) of the practice level x operation x block/session x strategy cell]. Obviously, this procedure was only needed in the ANOVAs on strategy efficiency and not in the ANOVAs on strategy selection. In Experiment 1, the number of cells replaced was 45 in the across-practice ANOVA and 10 in the post-practice ANOVA. In Experiment 2, 56 cells were replaced in the across-practice ANOVA and 34 in the post-practice ANOVA.

error rates (average = 3.1%). The main effect of block was significant, $F(3,33) = 10.10$. This effect consisted of a linear effect, $F(1,35) = 8.21$, which indicates that RTs decreased across practice blocks, and a quadratic effect, $F(1,35) = 25.89$, which indicates that RTs increased again in the test session. Block interacted with operation, $F(3,33) = 5.39$: RTs only decreased linearly for multiplication, $F(1,35) = 16.77$, but not for addition, $F < 1$. The insignificant block x strategy interaction, $F(3,33) = 1.42$ ($p = .25$), indicates that the decrease in RTs across practice blocks was significant for retrieval RTs, $F(1,35) = 36.85$ and for procedural RTs, $F(1,35) = 10.59$. The effect of practice was still present in the test session (i.e., first practice block vs. test session) for procedural RTs, $F(1,35) = 5.11$, but not for retrieval RTs, $F(1,35) = 1.80$ ($p = .19$). The nearly significant strategy x operation x block interaction, $F(3,35) = 2.47$ ($p = .07$), confirms that the linear decrease in procedural RTs across practice sessions was significant for multiplication, $F(1,35) = 20.12$ but not for addition, $F < 1$, whereas the linear decrease in retrieval RTs was significant in both operations but significantly larger in multiplication than in addition, $F(1,35) = 5.87$. No main or interaction effects with practice level were observed.

Post-practice effects. This section reports performance in the test session, i.e., after completion of the practice blocks. In the following analyses *all* problems (small and large ones) are included. To test practice effects on *strategy selection*, an ANOVA with problem size (small vs. large) as within-subjects variable and practice level (0x, 3x, 6x) and operation (addition vs. multiplication) as between-subjects variables was performed on percentages of retrieval use (see Table 3).

Table 2a
Retrieval response times (in msec) as a function of block[a], practice level, and operation. Standard errors are shown between brackets.

	Addition			Multiplication		
	3x practice	6x practice		3x practice	6x practice	
	Across blocks (=sessions)	Across blocks	Across sessions	Across blocks (=sessions)	Across blocks	Across sessions
1st practice block	945 (92)	937 (92)	1039 (84) 835 (74)	1378 (92)	1309 (97)	1392 (89) 1225 (78)
2nd practice block	789 (80)	889 (80)	835 (61) 891 (65)	1170 (80)	1106 (84)	1083 (65) 1129 (68)
3rd practice block	786 (69)	884 (69)	927 (68) 860 (56)	1108 (69)	1084 (73)	1123 (72) 1045 (59)
test session	957 (115)	923 (115)	923 (105)	1212 (115)	1296 (122)	1296 (111)

[a] For the 3x practice group, each practice block comprised one practice session. For the 6x practice group, each practice block comprised two practice sessions. For completeness, response times are provided for each single practice session in the 6x practiced participants as well.

Table 2b
Procedural response times (in msec) as a function of block [a], practice level, and operation. Standard errors are shown between brackets.

	Addition			Multiplication		
	3x practice	6x practice		3x practice	6x practice	
	Across blocks (=sessions)	Across blocks	Across sessions	Across blocks (=sessions)	Across blocks	Across sessions
1st practice block	1193 (400)	1211 (400)	1282 (238) 1140 (175)	2887 (400)	2933 (422)	3271 (222) 2594 (163)
2nd practice block	1103 (292)	1201 (292)	1157 (165) 1245 (139)	2530 (292)	2278 (308)	2331 (150) 2317 (117)
3rd practice block	1157 (212)	1226 (212)	1261 (134) 1101 (99)	2059 (212)	2085 (224)	1897 (130) 1920 (77)
test session	1115 (211)	1311 (211)	1311 (136)	1908 (211)	2456 (222)	2456 (144)

[a] For the 3x practice group, each practice block comprised one practice session. For the 6x practice group, each practice block comprised two practice sessions. For completeness, response times are provided for each single practice session in the 6x practiced participants as well.

Although there was no main effect of practice level, $F(2,53) = 2.11$ ($p = .13$), planned comparisons showed that 6x practiced participants used the retrieval strategy more often than 3x practiced participants did, $F(1,53) = 4.04$, whereas there was no difference between the 0x and 3x practiced participants, $F < 1$. Retrieval was used more frequently on small problems (77%) than on large problems (60%), $F(1,53) = 39.19$. Importantly, the difference in retrieval use between small and large problems was significantly higher for the 0x practiced participants (31%) than for the 3x practiced participants (15%) and the 6x practiced participants (4%), $F(1,53) = 5.89$ and $F(1,53) = 15.58$, respectively.

Table 3
Percentages of retrieval use in the test session, as a function of problem size, operation, and practice level. Standard errors are shown between brackets.

	Addition		Multiplication	
	Small	Large	Small	Large
0x practice	72 (7)	46 (9)	91 (7)	56 (9)
3x practice	56 (4)	39 (9)	82 (7)	69 (9)
6x practice	70 (7)	58 (9)	87 (7)	90 (10)

To test practice effects on *strategy efficiency*, an ANOVA on RTs was performed with practice level (0x, 3x, 6x) and operation (addition vs. multiplication) as between-subjects variables, and problem size (small vs. large) and strategy (retrieval vs. procedural) as within-subjects variables (see Table 4 and footnote 3). The main effect of practice level was significant, $F(2,53) = 6.07$. RTs were larger for the 0x practice level (1835 msec) than for both 3x and 6x practice levels (1300 msec and 1446 msec), $F(1,53) = 11.38$ and $F(1,53) = 5.85$, respectively. RTs did not differ between 3x and 6x practice levels, $F < 1$. Practice level interacted with strategy, $F(2,53) = 9.89$, and with operation, $F(2,53) = 4.43$. Practiced participants were more efficient in performing procedural strategies than unpracticed participants, $F(1,53) = 14.77$, whereas practiced and unpracticed participants were equally efficient in the retrieval strategy, $F < 1$. Furthermore, participants did not differ in efficiency on additions, $F < 1$, but practiced participants were more efficient than unpracticed participants on multiplications, $F(1,53) = 19.78$. The three-way interaction between practice level, strategy, and operation, $F(2,53) = 9.61$, confirmed that practiced participants were better than unpracticed participants in solving multiplication problems with procedural strategies, $F(1,53) = 27.79$, but not in solving multiplication problems via retrieval, $F < 1$, nor in solving addition problems with retrieval or procedural strategies (each $F < 1$).

There was no interaction between problem size and practice level, $F(2,53) = 1.69$ ($p = .19$). Planned comparisons confirmed that the retrieval problem-size effect (i.e., retrieval RTs large problems – retrieval RTs small problems) was significant in all groups and did not differ across groups, neither for addition nor for multiplication (each $F < 1$). The procedural problem-size effect (i.e., procedural RTs large problems – procedural RTs small problems) for multiplication was significantly larger for the 0x practice group than for the 3x practice group, $F(1,53) = 5.55$, but equally large in the 3x and 6x practice groups, $F(1,53) = 1.72$ ($p = .20$). The procedural problem-size effect for addition did not differ across groups either (each $F < 1$). These results indicate that the effect of practice on the problem-size effect originates from more efficient procedural strategy use; an effect that seems to be reliable for multiplication only.

Table 4
Retrieval and procedural response times (in msec) in the test session, as a function of problem size, operation, and practice level. Standard errors are shown between brackets.

Retrieval RTs	Addition		Multiplication	
	Small	Large	Small	Large
0x practice	889 (59)	986 (103)	1041 (59)	1300 (103)
3x practice	795 (59)	957 (103)	994 (59)	1212 (103)
6x practice	838 (59)	923 (103)	1025 (62)	1296 (109)
Procedural RTs	**Addition**		**Multiplication**	
	Small	Large	Small	Large
0x practice	1282 (327)	1303 (260)	3666 (327)	4213 (260)
3x practice	1146 (327)	1115 (260)	2272 (327)	1908 (260)
6x practice	1418 (327)	1311 (260)	2300 (344)	2457 (274)

As mentioned before, an additional test of *complex-arithmetic performance* (the French kit) was administered for each participant. Since this test was administered after the test session (and thus after the practice sessions for the practiced participants), scores on this test are indicative for transfer effects of simple-arithmetic practice on complex-arithmetic performance. An operation (addition or multiplication) x practice level (0x, 3x, 6x) ANOVA was run for each subtest of the French Kit (i.e., the addition subtest and the subtraction-multiplication subtest). For the addition subtest, no significant effects appeared (each $F < 1$), indicating no transfer effects at all. For the subtraction-multiplication subtest, in contrast, both main effects were significant. The practiced participants (3x and 6x) performed significantly better than the unpracticed (0x) participants, $F(1,53) = 5.1$ (means of 17.0, 16.7, and 13.8, respectively), indicating a transfer effect from simple to complex problems. Participants who had practiced multiplication problems performed significantly better than participants who had practiced addition problems, $F(1,55) = 5.5$ (means of 17.4 and 14.3, respectively), indicating that the transfer effect was operation specific.

Summary. Practice resulted in more frequent retrieval use, more efficient retrieval use, and more efficient procedural use. We may thus conclude that practice influenced both strategy selection and strategy efficiency. The effects were operation specific, though: selection effects were only present for addition whereas efficiency effects were only apparent for multiplication. Similarly, transfer effects from simple to complex problems were only significant for multiplication. The problem-size effect, finally, was reduced but did not completely disappear. Practice effects on the problem-size effect were associated with more frequent retrieval use and more efficient procedural use for large problems, but not with more efficient retrieval use for large problems.

All these interesting observations notwithstanding, this experiment had some drawbacks. First, there was no pre-practice session. Second, the strategy category "transformation" was very broad, and included many different strategies. Experiment 2 was meant to investigate the results obtained in Experiment 1 more thoroughly, by resolving the two problems above.

EXPERIMENT 2

The present experiment differed from Experiment 1 in four aspects. First, Experiment 2 included both a pre-practice and a post-practice session whereas Experiment 1 only included a post-practice session. In the pre-practice session, which was administered before the first practice session, participants had to solve all (i.e., both small and large) problems once. In the following practice sessions, only the large problems were practiced. In the post-practice session, all problems (small and large ones) had to be solved again. Second, more extensive strategy reports were required when participants used the transformation strategy. More specifically, when participants had chosen the transformation strategy, they had to answer an open question “How did you solve this problem? Describe extensively”. Their answer was written down by the experimenter. Afterwards, all answers have been put in different categories such as using 10 as reference point (e.g., 8 + 5 = 8 + 2 + 3; 9 x 6 = 10 x 6 – 6) or using a tie as reference point (e.g., 6 + 7 = 6 + 6 + 1; 6 x 7 = 6 x 6 + 6). Third, a speeded verification task was included. This task was administered twice: once before the pre-practice session and once after the post-practice session. This task was meant to test whether the practice effects on retrieval use observed in Experiment 1, were due to a real change in the retrieval network (i.e., the sensitivity) or to response biases. Indeed, trial-by-trial strategy reports have been criticized (e.g., Kirk & Ashcraft, 2001), since participants’ strategy reports may easily be biased by the experimenter’s suggestions. More specifically, participants might want to please the experimenter by reporting more frequent retrieval use without really using retrieval more frequently. As explained below, signal-detection theory can be used to disentangle ‘real’ practice effects on strategy selection from effects caused by response biases. Fourth, since Experiment 1 showed that 3 practice sessions were enough to obtain sensitive differences with the control condition, the number of practice sessions was restricted to 3. All participants thus took part in a simple-arithmetic task consisting of a pre-practice session, three practice sessions, and a post-practice session.

METHOD

Participants. Forty students (9 men and 31 women) at Ghent University participated in this Experiment. Half of them participated for course requirements and credits; the other half received €10 for participation. Their mean age was 20.0 years. None of them had participated in Experiment 1.

Procedure. Twenty participants completed simple addition problems and twenty participants completed simple multiplication problems. All participants also had to solve a complex-arithmetic test (the French kit) and to participate in a speeded verification task. The basic procedure of this second experiment was identical to the one used in Experiment 1, except the four changes described above. Hence, only the aspects of the procedure that were different for this experiment (e.g., the inclusion of a speeded verification task) are described underneath.

In the speeded verification task, participants had to verify simple additions or simple multiplications, depending on the operation they had to solve in the simple-arithmetic production task. Stimuli of the speeded verification task were presented in standard form (i.e., $a + b = c$ or a x $b = c$) in which a and b were one-digit numbers from 2 to 9. Half of the problems were presented with a correct solution whereas the other half were presented with an incorrect solution. The incorrect addition solutions were one or two units larger or smaller than the correct sum (e.g., 7 + 2 = 11). The incorrect multiplication solutions were 10% or 20% larger or smaller than the correct product (e.g., 3 x 4 = 10). To reduce interference effects, stimuli were excluded when (a) $c = a*b$ for addition problems (e.g., 3 + 2 = 6) or $c = a + b$ for multiplication problems (e.g., 2 x 3 = 5), (b) $c = a$ or $c = b$ (e.g., 2 + 2 = 2), (c) c = N*a or N*b for multiplication problems (e.g., 4 x 5 =16), and (d) c is even (uneven) while the correct solution is uneven (even) (e.g., 3 x 5 = 14).

The verification task consisted of 8 practice trials and 80 experimental trials. A trial started with a fixation point for 500 msec, after which the stimulus was presented until the participants responded or until the response deadline was met. The response deadlines were based on Experiment 1 and were calculated with the following formula: [*retrieval RT + 2 * standard deviation of retrieval RT*]. This measure was calculated separately for addition (1274 msec) and multiplication (1552 msec). After each practice trial, feedback was provided for one second, consisting of the word(s) “Correct” (when the answer was correct), “Incorrect” (when the answer was incorrect), or “Respond faster!” (when the participant’s response was slower than the response deadline). When the participant answered within the response deadline, his/her response time appeared on the screen as well. No feedback was provided in the experimental trials, although participants were strongly recommended to answer as fast and accurately as possible. All RTs higher than the response deadlines were discarded. The inter-trial interval was 500 msec. The speeded verification task was administered twice: one before practice and once after practice.

RESULTS

In the simple-arithmetic test, 722 trials (7.0%) were spoiled due to failures of the sound-activated relay. Since all the invalid trials met in the test session returned at the end of this session, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 53 (0.5%). Further, all incorrect trials (364 trials) and all trials on which participants selected the ‘Other’ category (38 trials) were deleted. All RTs more than 4 standard deviations from each participant’s mean (per operation) were discarded as outliers (76 trials). The results section is divided in two main parts. First, effects across all sessions are reported. Second, the pre-practice and post-practice sessions are compared to each other. The second part also includes the signal-detection analyses performed on the data obtained in the speeded verification task.

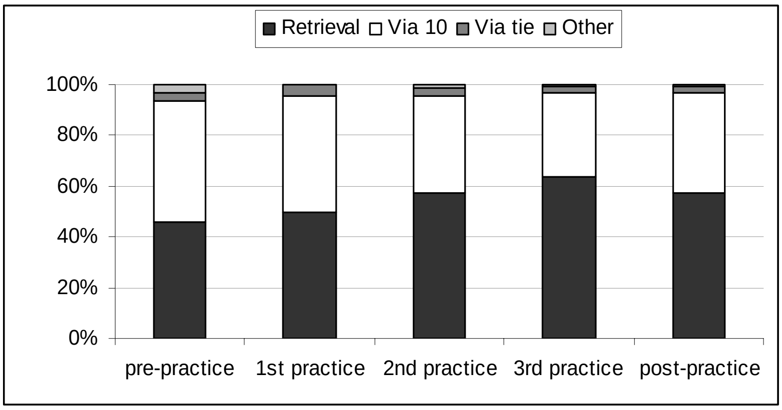
Effects during the practice sessions. Note that, as in Experiment 1, only the 12 most difficult problems are included in the following analyses, since only these problems were present in the three practice sessions and in the pre- and post-practice sessions. Figure 2 displays percentages usage of all strategies across all sessions, in order to present a global view on the effect of practice on *strategy selection*. Although retrieval was the most frequently used strategy, procedural strategies were used as well. Using 10 as a reference point was the most frequently chosen procedural strategy for both addition and multiplication. As in Experiment 1, the frequency of retrieval increased as practice proceeded, whereas frequency of procedural strategy use decreased.

To explicitly test these observations, an ANOVA was performed on percentages retrieval use, with session (pre-practice, first practice, second practice, third practice, post-practice) as within-subjects variable and operation (addition vs. multiplication) as between-subjects variable (see Table 5). The effect of session was significant both linearly and quadratically, $F(1,38) = 18.27$ and $F(1,38) = 17.70$, respectively. Reported use of direct retrieval increased linearly across the sessions, but it declined again in the post-practice session. The interaction between session and operation was not significant, $F(4,35) = 1.91$ ($p = .13$): the increase in retrieval use was significant both for addition and multiplication, $F(1,38) = 9.10$ and $F(1,18) = 9.16$, respectively.

Additional analyses were performed to test whether the increase in retrieval use was related to the initial amount of retrieval use. A median split on the percentage of retrieval use in the pre-practice session divided the participants in two groups. Twenty-one participants (7 in the addition group and 14 in the multiplication group) were labeled as frequent-retrieval users (86% retrieval use in the pre-practice session) and nineteen participants (13 in the addition group and 6 in the multiplication group) were labeled as infrequent-retrieval users (25% retrieval use in the pre-practice session). As in Experiment 1, the increase in retrieval use between the pre-practice session and the third practice session was larger for the infrequent-retrieval users (28%) than for the frequent retrieval users (11%), $t(38) = 2.69$. Anyhow, the decrease in retrieval use between the last practice session and the post-practice session did not differ across groups, $t < 1$ (5% for the frequent-retrieval users and 7% for the infrequent-retrieval users). Thus, between the pre-practice session and the post-practice session, the amount of retrieval use increased for both groups, 6% for the frequent-retrieval users, $t(20) = 2.60$ and 21% for the infrequent-retrieval users, $t(18) = 3.37$. The conclusion of these additional analyses runs parallel with the one in Experiment 1: practice effects are the largest in infrequent-retrieval users.

In order to test practice effects on *strategy efficiency*, an ANOVA on RTs was performed, with operation (addition vs. multiplication) as between-subjects variable, and session and strategy (retrieval vs. procedural) as within-subjects variables (see Tables 6a and 6b and footnote 3). The main effect of session indicated that RTs decreased linearly across sessions, $F(1,38) = 9.80$. Importantly, the linear decrease in RTs was significant for multiplication, $F(1,38) = 9.39$ but not for addition, $F(1,38) = 1.86$ ($p = .18$). More specifically, for addition, neither retrieval RTs nor procedural RTs decreased across sessions, $F(1,38) = 1.07$ and $F(1,38) = 1.19$, respectively. For multiplication, in contrast, both retrieval RTs and procedural RTs decreased across sessions, $F(1,38) = 9.46$ and $F(1,38) = 3.60$ ($p = .06$), respectively.

Addition



Multiplication

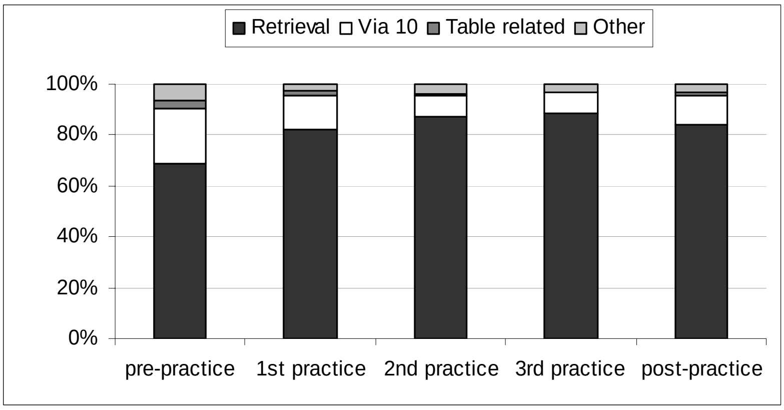


Figure 2
Percentage of use of different strategies across sessions.

Table 5
Percentages of retrieval use as a function of session and operation.
Standard errors are shown between brackets.

	Addition	Multiplication
Pre-practice	46 (7)	69 (7)
1st practice	51 (7)	82 (7)
2nd practice	57 (6)	88 (6)
3rd practice	64 (6)	89 (6)
Post-practice	57 (6)	84 (6)

Table 6a
Retrieval response times (in msec) as a function of session and operation.
Standard errors are shown between brackets.

Retrieval RTs	Addition	Multiplication
Pre-practice	904 (93)	1222 (93)
1st practice	807 (59)	1171 (59)
2nd practice	823 (54)	1007 (54)
3rd practice	766 (48)	986 (48)
Post-practice	847 (65)	1084 (65)

Table 6b
Procedural response times (in msec) as a function of session and operation.
Standard errors are shown between brackets.

Procedural RTs	Addition	Multiplication
Pre-practice	1182 (119)	1722 (119)
1st practice	1185 (97)	1698 (97)
2nd practice	1112 (102)	1741 (102)
3rd practice	1121 (113)	1549 (113)
Post-practice	1104 (83)	1606 (83)

Post-practice effects. In order to test practice effects on *strategy selection*, an ANOVA with problem size (small vs. large) and session (pre vs. post) as within-subjects variables, and operation (addition vs. multiplication) as between-subjects variable was performed on percentages of retrieval use on all problems (i.e., small and large ones) in the pre-practice and post-practice sessions (see Table 7). Percentages of retrieval use were higher in the post-practice session (77%) than in the pre-practice session (71%), $F(1,38) = 8.66$. Session interacted with problem size, though, $F(1,38) = 21.88$. Whereas retrieval use stayed equally high across the sessions for small problems, $F < 1$, it increased significantly for large problems, $F(1,38) = 15.13$. The difference in retrieval use between small and large problems was higher before practice (85% vs. 57%) than after practice (84% vs. 71%).

Table 7
Percentages of retrieval use as a function of session, problem size, and operation. Standard errors are shown between brackets.

	Addition		Multiplication	
	Small	Large	Small	Large
Pre-practice	77 (3)	46 (7)	94 (3)	69 (7)
Post-practice	74 (3)	57 (6)	94 (3)	84 (6)

Practice effects on *strategy efficiency* were tested with an ANOVA on RTs with problem size (small vs. large), session (pre vs. post) and strategy (retrieval vs. procedural) as within-subjects variables, and operation (addition vs. multiplication) as between-subjects variable (see Table 8 and footnote 3). RTs were faster in the post-practice session (1167 msec) than in the pre-practice session (1216 msec), $F(1,38) = 4.08$. This was true for both retrieval and procedural strategies, as appears from the insignificant session x strategy interaction, $F < 1$.

The main effect of problem size did not reach significance, $F(1,38) = 1.13$, but the interactions problem size x strategy and problem size x strategy x operation did, $F(1,38) = 25.93$ and $F(1,38) = 4.75$, respectively. The retrieval problem-size effect (i.e. retrieval RTs large problems – retrieval RTs small problems) was significant for both addition and multiplication, $F(1,38) = 15.65$ and $F(1,38) = 38.48$, respectively. Moreover, it did not change across sessions, $F < 1$ for addition and $F(1,38) = 1.64$ for multiplication. The procedural problem-size effect (i.e. procedural RTs large problems – procedural RTs small problems), in contrast, was not significant for addition, $F < 1$, but inversed for multiplication, $F(1,38) = 6.98$. More specifically, procedures were executed faster on large (i.e., practiced) multiplication problems than on small (i.e., unpracticed) multiplication problems. Obviously, this was only true in the post-practice session, $F(1,38) = 8.73$ and not in the pre-practice session, $F < 1$. In conclusion, practice enhanced the procedural efficiency of multiplication problems but not of addition problems.

Table 8
Retrieval and procedural response times (in msec) as a function of session, problem size, and operation. Standard errors are shown between brackets.

Retrieval RTs	Addition		Multiplication	
	Small	Large	Small	Large
Pre-practice	741 (45)	904 (93)	941 (45)	1222 (93)
Post-practice	706 (37)	847 (65)	889 (37)	1084 (65)
Procedural RTs	**Addition**		**Multiplication**	
	Small	Large	Small	Large
Pre-practice	1193 (137)	1182 (119)	1823 (137)	1722 (119)
Post-practice	1156 (146)	1104 (83)	1941 (146)	1606 (83)

As in Experiment 1, an additional test of *complex-arithmetic performance* (the French kit) was administered after the post-practice session. Separate ANOVAs were run for each subtest of the French Kit (i.e., the addition subtest and the subtraction-multiplication subtest) with operation (addition or multiplication) as the only independent variable. Participants having practiced simple additions scored slightly better on the complex-addition test than participants having practiced simple multiplications (15.6 vs. 14.9, respectively), but this effect did not reach significance, $F < 1$. The same was true for the complex-multiplication test, on which participants having practiced simple multiplications scored (insignificantly) better than participants having practiced simple additions (17.8 vs. 16.7, respectively), $F < 1$.

All *speeded-verification* trials on which participants had verified a correct addition or multiplication problem as correct were coded as hits. All trials on which participants had verified an incorrect addition or multiplication problem as correct were coded as false alarms. Using the software program of Van der Goten and Vandierendonck (1997), the signal-detection theory was used to determine the sensitivity (*d*') and the response bias (*c*) (see Table 9). A 2 x 2 ANOVA was conducted on these *d*' and *c* values, with operation (addition vs. multiplication) as between-subjects variable and session (pre vs. post) as within-subjects variable. The ANOVA on the sensitivity (*d'*) showed a significant main effect of operation, indicating that the sensitivity was higher for multiplication (3.23) than for addition (2.08), $F(1,38) = 24.91$. Moreover, the increase in sensitivity (pre vs. post practice) tended to be significant for multiplication, $F(1,38) = 3.67$ ($p = .06$) but not for addition, $F < 1$. The ANOVA on response bias (*c*) showed no significant effects (highest $F = 2.99$, $p = .10$). The practice effects on strategy selection could thus be attributed to real differences rather than to changes in response biases.

Table 9
Values of sensitivity (*d'*) and response bias (*c*) as a function of session and operation. Standard errors are shown between brackets.

	Sensitivity (*d'*)		Response bias (*c*)	
	Addition	Multiplication	Addition	Multiplication
Pre-practice	2.1 (0.2)	3.1 (0.2)	-0.1 (0.1)	-0.1 (0.1)
Post-practice	2.1 (0.2)	3.4 (0.2)	-0.1 (0.1)	0.0 (0.1)

Summary. As in Experiment 1, practice influenced both strategy selection and strategy efficiency. Indeed, direct memory retrieval was more frequent after practice than before, and retrieval and procedural use were more efficient after practice than before. In contrast to Experiment 1, where only additions were retrieved more frequently, practice increased the frequency of retrieval use for both addition and multiplication. Practice effects on strategy efficiency also showed a comparable effect as in Experiment 1, with larger practice effects on multiplication than on addition. The current experiment also confirmed that the problem-size effect was reduced by more frequent retrieval use and by more efficient procedural use, but not by more efficient retrieval use. The speeded-verification task finally, showed that the effects on strategy selection could not be attributed to response bias effects.

GENERAL DISCUSSION

The present study revealed some remarkable findings concerning the effects of practice on simple-arithmetic performance. First, practice enhanced frequency of retrieval use across practice sessions and thus influenced *strategy selection*. Second, as retrieval RTs decreased across practice sessions, *retrieval efficiency* was enhanced by practice as well. Third, practice enhanced *procedural efficiency*, since procedural RTs decreased across practice sessions. Fourth, participants who had been practicing simple-arithmetic problems were better in a complex-arithmetic test than were unpracticed participants (cf. Experiment 1), which indicates transfer from simple to complex-arithmetic problems. In the following, we aim at answering the four questions formulated in the introduction (i.e., practice effects on strategy selection, on strategy efficiency, on the problem-size effect, and transfer effects). We also check which arithmetic models are best fit to explain the results observed in this study.

PRACTICE EFFECTS ON STRATEGY SELECTION

Results of both experiments showed that retrieval use increased over the practice sessions, and was still used more often after practice than before. Practice effects on strategy selection were significant in both experiments and for both operations (addition and multiplication), with only one exception: in Experiment 1, retrieval use increased for addition problems only. As retrieval use was already reasonably high for multiplication problems before practice (i.e., 75%), there was little room left for amelioration (i.e., ceiling effect). Additional analyses indeed showed that practice effects on strategy selection were the strongest when the initial amount of retrieval use was rather low. It was more difficult to obtain practice effects on strategy selection as the level of experience increased.

However, the increase in reported retrieval use as a result of practice might be questioned, as it might be biased by demand effects (e.g., Kirk & Ashcraft, 2001). In that case, participants would (falsely) report more frequent retrieval use without really using retrieval more frequently. A signal detection analysis was used to disentangle effects caused by response bias and real practice effects. The results of this analysis suggest that the practice effects on retrieval frequency were likely not due to changes in the participants' response biases but to real sensitivity differences (i.e., changes in the memory network).

The increase in retrieval use as a result of practice can be explained by most theories. According to experience-based models, practice strengthens the association between a problem and its correct answer (Siegler & Shrager, 1984) or the association between a problem and the most efficient strategy (Siegler & Shipley, 1995). Stronger associations increase the possibility of direct memory retrieval. According to the instance theory of automatization (Logan, 1988), practice enhances retrieval efficiency but not procedural efficiency. As the race between retrieval and procedural strategies is always won by the fastest strategy, the race will be won by retrieval as practice progresses.

It should be noted, though, that retrieval frequency increased only slightly across the practice sessions and never reached 100% usage. Participants were quite rigid in their strategy choices and continued to use procedural strategies across the experiment. This observation cannot be accounted for by the models discussed above, as these models predict exclusive retrieval use after extensive practice. However, the amount of practice might have been insufficient in order to reach 100% retrieval use. The present study does not totally exclude the possibility that more intensive practice would result in exclusive retrieval use, as predicted by Siegler & Shipley (1995) and Logan (1988). Future research is needed to test whether very intensive practice would result in 100% retrieval use.

The reasonably high amount of procedural use after practice is in agreement with Baroody's (1983, 1984, 1985) theory, which states that people prefer procedures above retrieval because procedural knowledge is cognitively more economical than storing all individual facts in long-term memory. LeFevre et al. (1996a) also argue that procedural strategies are maintained, even after intensive practice. This can be explained by assuming that problem-procedure associations are as strong as (or even stronger than) problem-answer associations. This reasoning may also explain why changes in strategy selection are hard to make: when people perceive their commonly used procedural strategy as efficient, why should they switch to retrieval?

Another important observation concerning strategy selection was that retrieval use *decreased* when the test context changed. More specifically, retrieval use decreased when the practiced (large) problems were presented among the unpracticed (small) problems. This observation can be accounted for by Campbell's (1987a) network-interference theory of retrieval. Indeed, many answers to the unpracticed problems were false associates for the practiced problems. For example, 42 (the correct answer to the unpracticed problem 6 x 7) is a false associate for the practiced problem 6 x 8, and 40 (the correct answer to the unpracticed problem 5 x 8) is a false associate for the practiced problem 5 x 9. The exposure to these (unpracticed) problems in the test session resulted in the activation of answers that were less activated during the practice sessions. The activation of these answers, which are false associates of the practiced problems, produced interference: the activation of incorrect (but closely linked) answers of a problem interfered with the activation of the correct answer, resulting in less frequent retrieval use.

PRACTICE EFFECTS ON STRATEGY EFFICIENCY

Practice enhanced strategy efficiency. Both retrieval and procedural RTs decreased as a result of practice. However, this effect was longer-lasting for procedures than for retrieval. Practice effects on *retrieval* efficiency are predicted by the ASCM (Siegler & Shipley, 1995), the CMPL theory (Rickard, 1997), and the instance theory of automatization (Logan, 1988). Practice effects on *procedural* efficiency are harder to explain. Network theories of fact retrieval (e.g., Campbell, 1987a; Campbell & Graham, 1985) are silent about practice effects on procedural efficiency, and the instance theory of automatization (Logan, 1988) even precludes practice effects on procedural efficiency. The ASCM (Siegler & Shipley, 1995) and the theory of Baroody (1983, 1984, 1985), in contrast, can account for higher procedural efficiencies as a result of practice. The ASCM predicts an increase in the efficiency with which *each* strategy is executed, and thus predicts higher efficiencies in both retrieval and procedural strategies. According to Baroody, the development of arithmetic strategy use is rather due to a shift from slow procedural strategies to fast and automatic procedural strategies than to a shift from procedural strategies to retrieval strategies. Finally, the CMPL theory (Rickard, 1997) explains increases in procedural efficiency as a function of increases in retrieval efficiency. In this theory, a computation is first reduced to its simpler parts, each of which is resolved by direct memory retrieval (see also Anderson, 1993). Hence, both retrieval and procedural efficiency increase when memory traces are strengthened.

Furthermore, two interesting observations were made with regard to the increase in strategy efficiency as a result of practice. First, the increased strategy efficiencies decreased again in the test session, where both practiced (large) and unpracticed (small) problems had to be solved. Comparable results were obtained by Campbell (1987a), who observed a significant increase in RTs when practiced and unpracticed problems were tested all together, and by Rickard et al. (1994), who observed that multiplication and division RTs on an immediate test after practice sessions were reliably slower than expected by extrapolating. As already noted when discussing practice effects on strategy selection, such effects can be accounted for by Campbell's (1987a) network-interference theory of retrieval. The exposure to the unpracticed problems, of which the answers were false associates for the practiced problems, might have resulted in the activation of answers that were less active during the practice sessions. This activation of incorrect (but closely linked) answers interferes with the activation of the correct answer, resulting in slower RTs.

A second interesting observation with regard to strategy efficiency was that practice effects on strategy efficiency were much larger for multiplication than for addition, and especially when multiplications were solved by procedural strategies. Comparable effects have been observed in a previous study (Imbo, Vandierendonck, & Rosseel, in press f), in which arithmetic experience influenced strategy selection and strategy efficiency for multiplication problems only. There are several possible explanations for this effect. First, addition problems were solved faster than were multiplication problems. As can be seen in tables 2 and 6, initial retrieval and procedural RTs were slower for multiplication than for addition. Hence, multiplication efficiency was easier to increase than addition efficiency. Second, the procedures used to solve multiplication problems might be more consistent than those used to solve addition problems. Multiplication problems with a 9 (e.g., 9 x 7), for example, were consistently solved with the 'ten rule' (e.g., 9 x 7 = 10 x 7 – 7 = 70 – 7 = 63). Once this rule is sufficiently mastered, people are able to use this rule very efficiently (i.e., very fast and accurately). The available rules are less consistent for addition, and most of them involve counting, which is very time-consuming.

THE PROBLEM-SIZE EFFECT

According to Campbell and Xue (2001), there are as many as three sources for the problem-size effect: more frequent use of procedural strategies for large than for small problems, less efficient retrieval use for large than for small problems, and less efficient procedural use for large than for small problems. In the following, we discuss the effects of practice on all three sources.

First, practicing large problems influenced strategy selection, since retrieval was used more often on the practiced (i.e., large) problems as practice progressed. Consequently, the difference in retrieval use between large and small problems became smaller, reducing the problem-size effect. Second, practicing large problems did not change the retrieval problem-size effect. Retrieval was always slower for large problems than for small problems, and the difference in retrieval efficiency between large and small problems was not reduced by practice. Third, practicing large problems did change the procedural problem-size effect. Since procedures became faster for large problems as practice progressed, the procedural problem-size effect decreased. In Experiment 2, the decrease was so strong that the problem-size effect for multiplication inversed. We might thus conclude that more frequent retrieval use on large problems and more efficient procedural use on large problems reduced the problem-size effect as a result of practice. The difference in retrieval efficiency between small and large problems was not influenced by practice. Practice thus only influenced two out of three sources of the problem-size effect.

The retrieval problem-size effect was never eliminated (i.e., large problems were always solved more slowly than small problems), even at asymptotic response times, which is in line with previous research (e.g., Fendrich et al., 1993; Pauli et al., 1998). Several models however, predict that the problem-size effect should disappear with intensive practice if only retrieval strategies would be used. In experience-based models (e.g., Siegler & Shipley, 1995; Siegler & Shrager, 1984), for example, stronger problem-answer associations are created for small than for large problems. Since the length of RTs depends on the problem's association strength, answers are stated faster when problem-answer associations are stronger. However, each time an answer is given, the association between that answer and the problem increases – an increment that is twice as large for correct answers as for incorrect answers, but as large for small problems as for large problems. Practicing large problems should thus strengthen the problem-answer associations for these problems, resulting in equal retrieval RTs for small and large problems.

According to the instance theory of automatization (Logan, 1988), the retrieval problem-size effect should be eliminated as well. Practice creates automatization, which stands for the replacement of procedural strategies by direct memory retrieval. As long as large problems are encountered as frequently as necessary, they should be retrieved as fast as small problems.

In Campbell's (1995) network-interference model, in contrast, the problem-size effect is more robust against practice. In this model, the presentation of an arithmetic problem primes both the correct answer and its associated answers. Since large problems have more associated answers than small problems, interference is larger for large problems than for small problems. Since this explanation attributes the problem-size effect to structural aspects of memory representations, it should not be eliminated by practice (see also the interacting neighbors model of Verguts & Fias, 2005).

TRANSFER EFFECTS

The present results showed that calculation skill learning is very specific: Practice effects were limited to the operation that had been practiced. Practice effects did transfer over complexity though, as practiced participants were better than unpracticed participants in solving complex-arithmetic facts (cf. Experiment 1). It should be noted that transfer only occurred for multiplication but not for addition in Experiment 1, and was (insignificantly) larger for multiplication than for addition in Experiment 2.

The results thus indicate that transfer effects occur rather rarely. The data further suggested that transfer effects were limited to procedural strategies and absent for retrieval strategies. Indeed, procedures are applicable on several problems whereas retrieval is item specific. The data seem to confirm this statement. Indeed, (a) practice enhanced procedural efficiencies more strongly for multiplication than for addition (as discussed above), and (b) transfer occurred more clearly for multiplication than for addition. It can thus be argued that transfer effects are stronger for procedural strategies than for retrieval. Comparable effects (i.e., transfer for procedural strategies but not for retrieval strategies) have been observed in complex forms of mental arithmetic as well (e.g., Delazer et al., 2005). Our results also demonstrate that skilled arithmetic not only reflects efficient retrieval use of individual facts (declarative knowledge), but also efficient procedural strategy use (procedural knowledge). Indeed, when only retrieval use would be practiced, probably no transfer would occur.

The observation of transfer effects is in agreement with Baroody's theory (1983, 1984, 1985), which emphasizes the economy and all-round applicability of procedural knowledge. Indeed, procedural heuristics and principles can be more economically stored than many individual facts.

Moreover, as procedural knowledge is applicable to multiple problems, it can easily transfer to other problems.

Our data are also in agreement with the identical elements model of Rickard et al. (1994), although his theory is more retrieval-based (i.e., suited for simple-arithmetic performance *after* extended practice) than the theories of Baroody (1983, 1984, 1985), which are more procedure-based. The identical elements model assumes a distinct abstract representation for each unique combination of the basic elements (i.e., the operands and the required operation). Transfer is thus possible within the same operation but not between operations, which was observed in the present study. Rickard's (1997) CMPL theory is also consistent with transfer from simple to complex problems, provided that the retrievals involved in simple-arithmetic problem solving (e.g., 6 + 7) are also required in complex-arithmetic problem solving (e.g., 16 + 7).

Finally, the instance theory of automatization (Logan, 1988) predicts that transfer to novel stimuli should be poor to nonexistent. As automatization is item-based rather than process-based, it involves learning specific responses to specific stimuli.

CONCLUSION

Enhancing simple-arithmetic performance by practice is harder than it seems. Practiced participants used direct memory retrieval more often (i.e., a change in *strategy selection*), but tended to relapse in their old procedural strategies when the test context changed. Practice effects on *retrieval efficiency* were not overwhelming either, as they decreased after practice as well. When researchers (or tutors) want to conduct a practice or training program, they should try to avoid interference effects by constructing an overall item set (i.e., including all items) right from the beginning. Future research might also investigate whether a more intensive practice or training would create long-term effects on retrieval frequency or retrieval efficiency.

Practice effects on *procedural efficiency*, on the other hand, were large and persistent. This effect was operation dependent, however, as it was much stronger for multiplication than for addition. As procedural (but not retrieval) strategies can be applied to several problems, *transfer effects* were also stronger for multiplication than for addition. The current study also revealed some insights in the influence of practice on the *problem-size effect*. Whereas small problems were always retrieved faster than large ones, practice did reduce the problem-size effect by enhancing procedural efficiency on large problems. The problem-size effect was also reduced by strategy selection effects.

In reviewing several theories and models, we discovered that each model has its strengths and weaknesses. However, no model could explain all effects observed. The instance theory of automatization (Logan, 1988), was able to explain the replacement of procedural strategies by direct memory retrieval and the increase in retrieval efficiency. Practice effects on procedural efficiency and the concomitant transfer effects, in contrast, could not be explained by this theory. The latter two effects fit very well in the procedure-based theory of Baroody (1983, 1984, 1985), which is silent about increases in retrieval efficiency and retrieval frequency, though. The adaptive choice model of Siegler and Shipley (1995) and the CMPL theory of Rickard (1997) were able to explain practice effects on strategy selection, retrieval efficiency, and procedural efficiency. Finally, the network-interference theory of retrieval (Campbell, 1987a) was able to explain the decreasing performance when the test context changed.

In conclusion, there are several challenges for the arithmetic models. First, although most models (e.g., Campbell, 1987a; Campbell & Graham, 1985; Logan, 1988; Rickard, 1997; Siegler & Shipley, 1995) predict exclusive retrieval use after extensive practice, 100% retrieval use was never reached in the present study. Procedures were still used after practice. Moreover, changes in procedural efficiency were important sources of the decreasing problem-size effect. Arithmetic models should thus include parameters determining procedural characteristics (problem-procedure associations in addition to problem-answer associations). If procedural strategies are executed successfully, the problem-procedure association should be strengthened. In such models, procedural strategies should be automatically activated, without (or before) the answer is retrieved (e.g., Baroody, 1983, 1984, 1985; LeFevre et al., 1996a).

Second, practice never removed the retrieval problem-size effect: retrieving large problems took more time than retrieving small problems. This result is in agreement with the network-interference theory of retrieval (Campbell, 1995) but not with experience-based models (e.g., Siegler & Shipley, 1995; Siegler & Shrager, 1984) or the instance theory of automatization (Logan, 1988). Models should thus include both structural elements (e.g., problems that are associated with several answers which can interfere with each other) and experiential elements (e.g., the number of times a problem is encountered and/or solved).

Third, models should be able to explain what happens in the strategy selection process. According to the instance theory of automatization (Logan, 1988), both retrieval and procedural strategies are activated automatically, and the fastest one 'wins the race' and is selected. In the adaptive choice model of Siegler and Shipley (1995), there is a data base with information about the efficiency of each strategy, and this information is used in the strategy selection process. In this model, only one strategy can be executed at one time. The CMPL theory of Rickard (1997) also excludes parallel completion of retrieval and nonretrieval strategies. The procedural theory of Baroody (1983, 1984, 1985) is silent about what really happens in the strategy selection process, as are retrieval models (Campbell, 1987a; Campbell & Graham, 1985), which take for granted that retrieval is the only strategy available in adults.

CHAPTER 4
THE ROLE OF PHONOLOGICAL AND EXECUTIVE WORKING-MEMORY RESOURCES IN SIMPLE-ARITHMETIC STRATEGIES

European Journal of Cognitive Psychology (in press)[1,2]

The current study investigated the role of the central executive and the phonological loop in arithmetic strategies to solve simple addition problems (Experiment 1) and simple subtraction problems (Experiment 2). The choice/no-choice method was used to investigate strategy execution and strategy selection independently. The central executive was involved in both retrieval and procedural strategies, but played a larger role in the latter than in the former. Active phonological processes played a role in procedural strategies only. Finally, passive phonological resources were only needed when counting was used to solve subtraction problems. No effects of working-memory load on strategy selection were observed.

---

[1] This paper was co-authored by André Vandierendonck.

[2] Thanks are extended to Koen Luwel, Jeff Bisanz, and two anonymous reviewers for their helpful comments on previous drafts of this article.

INTRODUCTION

Being able to solve arithmetic problems mentally (i.e., without using a calculator or a similar device) is a skill which is very useful in daily life. During the past decade, many studies have shown that mental arithmetic relies – among other things – on a well functioning working memory (see DeStefano & LeFevre, 2004, for review). Although working-memory resources might fulfill a role in several subprocesses of the arithmetic problem-solving process (e.g., problem encoding, accessing and searching long-term memory, calculating the correct answer, stating the answer), the current study concentrates on the role of working memory in the processing stages that take place *after* the problem has been encoded and *before* the answer is stated.

In these specific processing stages, people might use a variety of strategies to solve the arithmetic problem (e.g., Hecht, 1999; LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). They might retrieve the answer directly from long-term memory, or they might use a nonretrieval (procedural) strategy to solve the problem. Two frequently used procedural strategies are transformation, in which the problem is solved by referring to related operations or by deriving the answer from known facts, and counting, in which participants count one-by-one to reach the correct answer. However, not much is known about the role of working memory across these different arithmetic strategies. More specifically, it is not known whether the execution of *all* arithmetic strategies does rely on working-memory resources. It is not known either whether or not *all* working-memory components are needed across the different arithmetic strategies.

Working memory, as proposed in the model of Baddeley and Hitch (1974), indeed consists of several components: a central executive and two slave systems[3]. The central executive can be seen as a system with limited capacity that allocates attentional resources to various processes, such as controlling, planning, sequencing, and switching activities. This component also integrates and coordinates the activities of the slave systems, the phonological loop and the visuo-spatial sketchpad. The phonological loop maintains and manipulates verbal-phonological information whereas the visuo-spatial sketchpad maintains and manipulates visuo-spatial information. The phonological loop further consists of two components: an active subvocal rehearsal process and a passive, phonologically based store.

Previous studies showed that the central executive is always needed to solve simple-arithmetic problems (De Rammelaere, Stuyven, & Vandierendonck, 1999, 2001; De Rammelaere & Vandierendonck, 2001; Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Muyllaert, 2006; Deschuyteneer, Vandierendonck, & Coeman, 2007; Hecht, 2002; Lemaire, Abdi, & Fayol, 1996; Seitz & Schumann-Hengsteler, 2000, 2002; Seyler, Kirk, & Ashcraft, 2003). Note that simple-arithmetic problems encompass all problems with correct answers up until 20 (e.g., 4 + 8, 13 - 6), as opposed to complex-arithmetic problems, which encompass more multi-digit problems (e.g., 36 + 72, 125 - 46). The role of the phonological loop in simple arithmetic is less clear, however. DeStefano & LeFevre (2004) note that the role of this working-memory component may depend on several factors, such as educational experience and the operation studied. Indeed, most studies with western participants did not observe a significant role of the phonological loop in solving simple addition or multiplication problems (e.g., De Rammelaere et al., 1999, 2001; Seitz & Schumann-Hengsteler, 2000; but see Lemaire et al., 1996), whereas at least one study with East-Asian participants did observe a

[3] More recently, a third slave system was proposed by Baddeley (2000), namely the episodic buffer. This system integrates information in both other slave systems with information from long-term memory.

significant role of the phonological loop in solving simple multiplication problems but not in solving simple subtraction problems (Lee & Kang, 2002).

What has been neglected in the aforementioned studies, however, is the fact that people use several strategies to solve simple-arithmetic problems (e.g., Hecht, 1999; LeFevre et al., 1996a, 1996b). Consequently, the issue of working-memory involvement across different arithmetical strategies has scarcely been investigated. To our knowledge, only Hecht (2002) and Seyler et al. (2003) have published research on this topic.

In Hecht (2002), participants had to verify simple addition equations (e.g., 4 + 3 = 8, true/false?) under no-load conditions and conditions in which the central executive or phonological loop were loaded. After each trial, participants had to report which strategy they had used. As the pattern of strategy selection (i.e., percentages of chosen strategies) was comparable between no-load and working-memory load conditions, Hecht concludes that phonological or executive working-memory loads do not influence strategy selection. Strategy efficiency (i.e., solution times of the strategies), in contrast, was impaired by reduced availability of working-memory resources. More precisely, Hecht observed that all strategies (i.e., retrieval, transformation, and counting) were slowed down under executive working-memory loads, whereas only the counting strategy was slowed down under phonological working-memory loads. Based on regression analyses, however, Hecht concludes that retrieval does not rely on the central executive, whereas the counting strategy would rely on both the central executive and the phonological loop.

In Seyler et al. (2003), participants had to solve simple subtraction problems (e.g., 8 – 4 = ?) while their working memory was loaded by means of a task in which 2-, 4-, or 6-letter strings had to be remembered. As subtraction performance was slower and more erroneous when participants' working memory was loaded, Seyler et al. conclude that the processing of subtraction facts relies heavily on working memory. Otherwise stated, strategy efficiency decreased under working-memory load, and this was especially the case for participants with low working-memory spans. Moreover, Seyler et al. observed that working memory was more involved in procedural strategies than in direct memory retrieval. Although Seyler et al. (2003) do not report specific data about the secondary task they used, it may be assumed that the task was primarily loading the phonological loop, and to a lesser extent the central executive.

Based on the studies by Hecht (2002) and Seyler et al. (2003), one could conclude that executive and phonological working-memory components are used in procedural strategies but not in direct memory retrieval. However, it is difficult to draw strong conclusions from these studies since (a) Hecht used an addition verification task (e.g., 8 + 5 = 12, true/false?) whereas Seyler et al. used a subtraction production task (e.g., 12 – 5 = ?), (b) Hecht loaded working memory phonologically and executively whereas it is unclear which working-memory components were loaded by the secondary task used by Seyler et al., and (c) in both Hecht's and Seyler et al.'s study participants were always free to choose the strategy they wanted, which may have biased strategy efficiency data, as explained further in this article.

Although Hecht (2002) and Seyler et al. (2003) already addressed the role of working memory in simple-arithmetic strategies, the current study was designed to achieve additional insight. First, in the current study, both addition and subtraction problems on which participants had to produce the correct answer themselves (i.e., production tasks) were used. Verification strategies indeed differ from strategies used in production tasks (e.g., Campbell & Tarling, 1996; Krueger & Hallford, 1984; Lemaire & Fayol, 1995; Zbrodoff & Logan, 1990). The face validity is also larger in production tasks than in verification tasks, as they are more frequently used in daily life. Moreover, solving simple subtraction problems in adults received little attention up until now (but see Barrouillet & Fayol, 1998;

Campbell & Xue, 2001; Deschuyteneer et al., 2007; Geary, Frensch, & Wiley, 1993b; Seyler et al., 2003).

Second, we wanted to clarify the role of both executive and phonological working-memory components in simple-arithmetic strategies. As the role of the visuo-spatial sketchpad in mental arithmetic is still unclear (e.g., DeStefano & LeFevre, 2004), it was decided not to include this working-memory component in the current project. The phonological loop was further subdivided in its two components (the active rehearsal process and the passive phonological store), and the role of both components was investigated. More specifically, retaining a 3- or 5-letter string in memory was used to load the active rehearsal process, whereas irrelevant speech was used to load the passive phonological store. Salamé and Baddeley (1982) indeed showed that the passive phonological store is accessed directly by speech while it leaves the active rehearsal process unaffected. It should be noted, however, that tasks loading the active rehearsal process rely on the passive phonological store as well. Finally, a continuous choice reaction time task (CRT task) was used to load the central executive. Szmalec, Vandierendonck, and Kemps (2005) have shown that this task interferes with the central executive, while the load on the slave systems is negligible. The CRT task has already been fruitfully adopted in mental-arithmetic studies (e.g., Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer et al., 2006, 2007).

Third, we aimed not only at investigating the role of working memory in strategy efficiency (i.e., how fast are strategies executed?) but also at investigating the role of working memory in strategy selection (i.e., which strategies do people choose?). The choice/no-choice method was used to investigate both strategy components (efficiency and selection) independently. As convincingly argued by Siegler and Lemaire (1997), strategy efficiency data obtained in choice conditions might be biased by selection effects. This might have been the case in the studies of Hecht (2002) and Seyler et al. (2003), since these studies only involved a choice condition in which participants were free to choose the strategy they wanted. In the choice/no-choice method, however, each participant is tested under two types of conditions. In the choice condition, participants are free to choose which strategy they want to solve the arithmetic problems. In the no-choice conditions, participants are forced to solve all the problems with one particular strategy. This obligatory use of one particular strategy on all problems precludes selective assignments of strategies to problems and thus yields unbiased strategy efficiency data. There are as many no-choice conditions as there are strategies available in the choice condition. Data obtained in no-choice conditions provide information about strategy efficiency, whereas data gathered in the choice condition provide information about strategy selection. The choice/no-choice method has already been used with arithmetic problems, both in children (e.g., Imbo & Vandierendonck, in press d; Lemaire & Lecacheur, 2002; Torbeyns, Verschaffel, & Ghesquière, 2002, 2004a, 2004b, 2005a, 2005b) and in young and older adults (e.g., Imbo, Duverne, & Lemaire, in press a; Siegler & Lemaire, 1997).

Concerning strategy efficiency, it was expected that the central executive would play a role in all strategies, but to a larger extent in procedural strategies than in retrieval. Indeed, several processes which are supposed to rely on executive working-memory resources (e.g., the manipulation and calculation of digits) are needed in procedural strategies but not in direct memory retrieval. Anyhow, as accessing long-term memory and selecting the correct answer are processes which might rely on the central executive, effects of an executive working-memory load on direct memory retrieval were expected as well. Because temporarily storing intermediate results is only needed in procedural and not in retrieval strategies, it was predicted that the phonological working-memory components would play a role in the procedural strategies but not in retrieval. As executive resources fulfill coordination and manipulation functions whereas phonological resources only fulfill storage functions, we expected that the role of the central executive would be larger than the role of phonological working-memory resources. As efficient arithmetic performance not only requests passive storage but also active maintenance of partial results, we further expected that the role of the active phonological rehearsal process would be larger than the role of the passive phonological store. Concerning strategy selection, finally, no effects of working-memory load were expected. This hypothesis was based on previous research which did not find load effects on strategy selection either (e.g., Hecht, 2002).

EXPERIMENT 1: ADDITION

METHOD

Participants. Forty-five first-year psychology students (5 men and 40 women) at Ghent University participated for course requirements and credits. Their mean age was 20 years and 0 months.

Stimuli. Stimuli of the primary task (i.e., the simple-arithmetic task) consisted of simple addition problems that were composed of pairs of numbers between 2 and 9. Problems involving 0 or 1 as an operand or answer (e.g., 5 + 0) and tie problems (e.g., 3 + 3) were excluded. All problems crossed 10 (e.g., 3 + 8). Since commuted pairs (e.g., 9 + 4 and 4 + 9) were considered as two different problems, this resulted in 32 addition problems (ranging from 2 + 9 to 9 + 8). Stimuli of the executive secondary task (i.e., the CRT task) consisted of low tones (262 Hz) and high tones (524 Hz) that were sequentially presented with an interval of 900 or 1500 msec. Participants had to press the 4 on the numerical keyboard when they heard a high tone and the 1 when a low tone was presented. The duration of each tone was 200 msec.

Two tasks were used to load the active phonological rehearsal process. Doing so, we wanted to differentiate between the ‘easier’ and ‘more difficult’ tasks used in the past. Indeed, the phonological secondary tasks used in previous studies strongly differed from each other (DeStefano & LeFevre, 2004). The main difference across both active phonological tasks is the amount of letters that has to be maintained. In the current study, stimuli of the easier task consisted of letter strings of 3 consonants (e.g., T K X) whereas stimuli of the more difficult task consisted of letter strings of 5 consonants (e.g., F S W R M). These consonants were read aloud by the experimenter. The participant had to retain these letters. After three simple-arithmetic problems, participants in the 3-letter task had to repeat the letters in the correct order. Participants in the 5-letter task had to decide whether the order of two adjacent letters that were read aloud by the experimenter was correct (e.g., S W) or incorrect (e.g., W S). The replacement of letter repetition by order verification was based on pilot studies which had showed that repeating all 5 letters in the 5-letter task was too demanding. Replacing letter repetition by order verification made the 5-letter task easier. However, because retaining 5 letters in memory is more demanding than retaining only 3 letters in memory; the 5-letter task was still more difficult than the 3-letter task. Being more difficult, it is possible that the 5-letter task would also demand executive working-memory resources. The results might give a decisive answer about this issue, and will be discussed further in this paper. For both active phonological tasks, a new 3- or 5-letter string was presented by the experimenter following the response of the participant.

The passive phonological task (irrelevant speech) consisted of dialogues between several people in the Swedish language, which were taken from a compact disc used in language courses. The Swedish dialogues were presented with an acceptable loudness (i.e., around 70 dB) through the headphones. None of the participants had any notion of Swedish.

Procedure. Each participant was tested individually in a quiet room for approximately 50 minutes. The experiment was started with short questions about the age of the participant, his/her study curriculum (i.e., the number of weekly mathematics lessons during the last year of secondary school), and calculator use (i.e., on a rating scale from 1 “never” to 5 "always"). All participants solved the simple-arithmetic problems in two sessions: one in which no working-memory component was loaded, and one in which one working-memory component (the central executive, the passive phonological store, or the active phonological rehearsal process) was loaded. The working-memory load differed across participants: for 10 participants the central executive was loaded, for 10 participants the passive phonological store was loaded, for 15 participants the active phonological rehearsal process was loaded with the 5-letter task, and for 10 participants the active phonological rehearsal process was loaded with the 3-letter task.

For the executive secondary task and the active phonological tasks, single-task data were obtained as well. To this end, participants had to carry out the secondary task for 2 minutes in absence of the primary task. An interval of 15 seconds was used between the 3-letter string and the question to repeat (in the 3-letter task) and between the 5-letter string and the 2-letter probe (in the 5-letter task). The secondary-task-only execution took place just before the execution of the primary task in combination with the respective secondary task. This permitted the participants to get used to the secondary-task execution.

Both no-load and load sessions consisted of four conditions: first the choice condition[4], and then three no-choice conditions, the order of which was randomized across participants. The choice condition started with comprehensive explanations about the simple-arithmetic task and the

---

[4] In both no-load sessions and load sessions, choice conditions were administered first in order to exclude influence of no-choice conditions on the choice condition. However, as there were two choice conditions (one in the no-load session and one in the load session), order effects still might have occurred. A paired-samples *t*-test indicated a small but significant difference between 1st session (no-load or load) choice RTs and 2nd session (load or no-load) choice RTs, $t(84) = 2.3$, with RTs in the 2nd session being 70 msec smaller than than in the 1st one. However, as working-memory load was counterbalanced across participants, the bias resulting from this general speeding effect should be small to non-existent.

strategy reportage. In the choice condition, 6 practice problems and 32 experimental problems were presented. Subsequently, explanations about the no-choice conditions were given, and participants had to solve 32 simple-arithmetic problems in each of the three no-choice conditions. After a break of approximately 5 minutes, the second session was administered. This session also consisted of one choice condition and three no-choice conditions. The participants who were enrolled in a dual-task session first, now solved the simple-arithmetic problems without secondary task, whereas this order was reversed for the other half of the participants.

A trial started with a fixation point for 500 msec. Then the arithmetic problem appeared in the center of the screen. The problems were presented horizontally in Arabic format, with the operation sign at the fixation point. The problem remained on screen until the participant responded. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, participants wore a microphone which was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 msec. On each trial, feedback was presented to the participants, a green 'Correct' when their answer was correct, and a red 'Fout' when it was not.

In the choice condition, participants were free to choose the strategy they wanted. Trial-by-trial self reports were used to know which strategy the participants had used. Immediately after solving each problem, they had to report verbally which of the four strategies displayed on the screen they had used (see also Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler et al., 2003): *(1) Retrieval: You solve the problem by remembering or knowing the answer directly from memory; (2) Count: You solve the problem by counting a certain number of times to get the answer; (3) Transform: You solve the problem by referring to related operations or by deriving the answer from known facts; (4) Other: You solve the problem by a strategy unlisted here, or you do not know what* *strategy that you used to solve the problem.* These four choices had been extensively explained by the experimenter, and it was emphasized that the presented strategies were not meant to encourage use of a particular strategy.

In the no-choice conditions, participants were requested to use one particular strategy to solve all problems. In no-choice/retrieval, they were asked to retrieve the answer. More specifically, they had to say the answer that first popped into their head. In no-choice/transform, they were asked to transform the problem by making an intermediate step to 10 (e.g., 8 + 5 = 8 + 2 + 3 = 10 + 3 = 13). In no-choice/count, finally, they had to count one-by-one (subvocally) until they reached the correct total (e.g., 7 + 4 = 8… 9… 10… 11). After having solved the problem, participants also had to answer with ‘yes’ or ‘no’ whether they had succeeded in using the forced strategy. The answer of the participant, the strategy information, and the validity of the trial were recorded online by the experimenter. All invalid trials (e.g., failures of the voice-activated relay) were discarded and returned at the end of the block, which minimized data loss due to unwanted failures.

After the simple-arithmetic experiment, all participants completed a pen-and-paper test of complex arithmetic, the French Kit (French, Ekstrom, & Price, 1963). The test consisted of two subtests, one page with complex addition problems and one page with complex subtraction and multiplication problems. Participants were given 2 minutes per page, and were instructed to solve the problems as fast and accurately as possible. The number of correct answers on both subtests were summed to yield a total score.

RESULTS

Of all trials 7.47% was spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 1.28%. Further, all incorrect trials (2.41%), all choice trials on which participants reported having used a strategy ‘Other’

(0.05%), and all no-choice trials on which participants failed to use the forced strategy (11.80%) were deleted. All data were analyzed on the basis of the multivariate general linear model; and all reported results are considered to be significant if $p < .05$, unless mentioned otherwise. To test whether the four subject groups (i.e., loaded by the passive phonological task, the active phonological task with 3 letters, the active phonological task with 5 letters, or the executive task) differed from each other, several univariate analyses of variance (ANOVAs) were conducted. A first ANOVA was conducted on the scores on the French Kit and showed that the four groups did not differ in mathematical skill, $F < 1$ (means of 28.4, 30.3, 27.8, and 31.9, respectively). A second ANOVA, conducted on the scores of the calculator-use questionnaire, showed that the four groups did not differ in their reported calculator use, $F < 1$ (means of 3.8, 3.4, 3.5, and 3.6). The last ANOVA, conducted on the amount of arithmetic lessons in the last year of secondary school, showed no group differences either, $F < 1$ (means of 3.8, 3.9, 4.5, and 4.6).

Strategy efficiency. Only the RTs uncontaminated by strategy choices (i.e., no-choice RTs) will be considered, since only these RTs provide clear data concerning strategy efficiency. A 4 x 2 x 3 ANOVA was conducted on correct RTs with working-memory component (passive phonological, active phonological with 3 letters, active phonological with 5 letters, or executive) as between-subjects effect, and load (no load vs. load) and strategy (retrieval, transformation, counting) as within-subjects effects (see Table 1). The main effect of load was significant. RTs were higher under load than under no-load, $F(1,41) = 21.72$, $MSe = 83256$. The main effect of strategy was significant as well, $F(2,40) = 175.58$, $MSe = 457279$, and indicated that RTs differed significantly across strategies, with retrieval being faster than transformation, $F(1,41) = 148.64$, $MSe = 80048$, which was in its turn faster than counting, $F(1,41) = 142.05$, $MSe = 649757$. Load and strategy interacted, $F(2,40) = 5.60$, $MSe = 232424$. Although the effect of load was highly significant for each single strategy, it was larger on counting and transformation than on retrieval, $F(1,41) = 6.44$ and $F(1,41) = 8.65$, respectively. The effect of load did not differ between transformation and counting, $F(1,41) = 1.82$ ($p = .18$). Although the effect of working-memory component did not reach significance ($F < 1$), there was a significant interaction between working-memory component and load, $F(3,41) = 6.89$. This interaction showed that the effect of working-memory load was significant for the active phonological component as measured by the 5-letter task, $F(1,41) = 40.21$, and for the executive component, $F(1,41) = 9.73$, but not for the passive phonological component ($F < 1$) or the active phonological component as measured by the 3-letter task ($F < 1$).

Table 1 (continued on next page)
No-choice RTs (in msec) as a function of load, working-memory component[a], and strategy. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Retrieval	No load	849 (81)	915 (81)	879 (66)	809 (81)	863 (39)
	Load	840 (83)	851 (83)	1074 (67)	940 (83)	926 (40)
	Mean	845 (79)	883 (79)	976 (64)	874 (79)	894 (38)

[a] PHON = phonological, CE = central executive.

Table 1
No-choice RTs (in msec) as a function of load, working-memory component[a], and strategy. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Transformation	No load	1388 (120)	1365 (120)	1400 (98)	1183 (120)	1334 (57)
	Load	1357 (149)	1458 (149)	1704 (121)	1479 (149)	1499 (71)
	Mean	1373 (131)	1412 (131)	1552 (107)	1331 (131)	1417 (63)
Counting	No load	2811 (223)	2854 (223)	2572 (182)	2707 (223)	2736 (107)
	Load	2922 (326)	2894 (326)	3231 (266)	2976 (326)	3006 (156)
	Mean	2867 (265)	2874 (265)	2902 (216)	2842 (265)	2871 (127)
All strategies	No load	1682 (109)	1711 (109)	1617 (89)	1566 (109)	1644 (52)
	Load	1706 (151)	1735 (151)	2003 (123)	1789 (151)	1811 (72)
	Mean	1695 (126)	1723 (126)	1810 (103)	1682 (126)	1727 (60)

[a] PHON = phonological, CE = central executive.

This interpretation was confirmed with separate ANOVAs for each single strategy, which tested the effects of the different working-memory loads. The effects of passive phonological loads and active phonological loads (as measured by the 3-letter task) were negligible for retrieval, transformation, and counting strategies. The effects of executive loads and active phonological loads (as measured with the 5-letter task) were highly significant for all strategies. The ANOVA on all strategies confirmed that, although an executive working-memory load affected all strategy RTs, the effect was smaller in retrieval than in transformation, $F(1,41) = 5.20$, but did not differ between counting and transformation, $F < 1$. Similarly, although an active phonological load (as measured by the 5-letter task) affected all strategy RTs, the effect tended to be smaller in retrieval than in transformation, $F(1,41) = 3.41$ ($p = .07$) and was smaller in transformation than in counting, $F(1,41) = 7.22$.

Strategy selection. In order to investigate effects on strategy selection, a 4 x 2 x 3 ANOVA was conducted on percentages strategy use (in the choice condition), with working-memory component as between-subjects effect, and load and strategy as within-subjects effects (see Table 2). All three strategies were used spontaneously by the participants, but the main effect of strategy, $F(2,40) = 121.35$, $MSe = 1488$, indicated that the percentage of use varied across strategies. Retrieval (51%) and transformation (44%) were used more frequently than counting (5%), $F(1,41) = 95.76$, $MSe = 950$ and $F(1,41) = 67.69$, $MSe = 953$, respectively. There was no difference between the percentage retrieval use and the percentage transformation use, $F < 1$. The main effects of load and working-memory component did not reach significance, and neither did any interaction (highest $F = 1.4$).

Table 2
Strategy use in the choice condition (in %) as a function of load and working-memory component[a] for all three strategies. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Retrieval	No load	71 (8)	44 (8)	53 (7)	38 (8)	52 (4)
	Load	64 (9)	36 (9)	63 (8)	39 (9)	50 (4)
Transformation	No load	25 (8)	53 (8)	39 (7)	53 (8)	42 (4)
	Load	32 (9)	58 (9)	33 (8)	56 (9)	45 (4)
Counting	No load	4 (5)	3 (5)	9 (4)	9 (5)	6 (2)
	Load	4 (3)	5 (3)	5 (3)	5 (3)	5 (2)

[a] PHON = phonological, CE = central executive.

Secondary task performance. Secondary task performance can be found in Table 3. Performance on the CRT task was significantly faster and more accurate when executed alone than when executed simultaneously with the primary task (choice and no-choice conditions taken together), $F(1,9) = 81.47$, $MSe = 1492$ and $F(1,9) = 99.98$, $MSe = 187$, respectively. CRT speed was also lower in no-choice/retrieval and no-choice/count than in choice conditions, $F(1,9) = 5.54$, $MSe = 1886$ and $F(1,9) = 5.35$, $MSe = 2323$, respectively. CRT accuracy did not differ across the choice condition and the three no-choice conditions. When few executive working-memory resources are left, performance was thus impaired not only on the primary task but also on the secondary task (cf. Hegarty, Shah, & Miyake, 2000).

The performance on the active phonological tasks with 3 and 5 letters was more accurate when executed alone than when executed simultaneously with the primary task (choice and no-choice conditions taken together), $F(1,9) = 14.92$, $MSe = 256$ and $F(1,14) = 26.50$, $MSe = 52$, respectively. Performance on the 3-letter task did not differ across choice and no-choice conditions, but performance on the 5-letter task was more accurate in the no-choice/count condition than in choice and no-choice/retrieval conditions, $F(1,14) = 13.38$, $MSe = 55$ and $F(1,14) = 11.17$, $MSe = 52$, respectively. The other comparisons did not reach significance.

SUMMARY

The analyses on strategy efficiency showed that not all working-memory loads affected the strategies needed to solve simple addition problems. More specifically, performance was affected by an executive working-memory load and an active phonological working-memory load (as measured by the 5-letter task) but not by a passive phonological working-memory load. Further analyses showed that this assertion accounted for all three strategies. Thus: retrieval, transformation and counting RTs all increased under an executive load and under an active phonological load (i.e., the 5-letter task), but not under a passive phonological load. However, procedural strategies were more affected by executive and phonological loads than retrieval strategies were. The analyses on strategy selection showed that retrieval was the most frequently used strategy, followed by transformation. Counting was used rather rarely. There was no effect of working-memory load on percentage strategy use. The next experiment, which was similar in design as Experiment 1, investigated the role of the different working-memory components in subtraction problems.

Table 3
RTs (in msec) and accuracies (in %) on the CRT task and accuracies (in %) on the active phonological tasks for single-task, choice, and no-choice conditions.
Standard deviations are shown between brackets.

Experiment 1	RT CRT task	ACC CRT task	ACC 3-letter task	ACC 5-letter task
Single	503 (131)	95 (8)	87 (20)	92 (6)
Choice	597 (325)	48 (14)	70 (16)	77 (10)
No-choice/retrieval	643 (355)	46 (15)	69 (20)	78 (13)
No-choice/transform	618 (347)	49 (17)	70 (21)	82 (13)
No-choice/count	647 (339)	45 (22)	76 (12)	87 (8)
Experiment 2	**RT CRT task**	**ACC CRT task**	**ACC 3-letter task**	**ACC 5-letter task**
Single	516 (131)	98 (3)	98 (5)	88 (13)
Choice	644 (391)	46 (9)	65 (22)	71 (7)
No-choice/retrieval	596 (360)	49 (15)	73 (25)	82 (17)
No-choice/transform	587 (352)	49 (13)	84 (17)	75 (8)
No-choice/count	624 (390)	42 (12)	72 (14)	69 (21)

EXPERIMENT 2: SUBTRACTION

METHOD

Participants. Forty first-year psychology students (10 men and 30 women) at Ghent University participated for course requirements and credits. Their mean age was 19 years and 4 months. None of them had participated in Experiment 1. There were 10 participants in each working-memory load condition.

Stimuli and Procedure. The 32 subtraction problems were the reverse of the addition problems used in Experiment 1, and thus crossed 10 as well (e.g., 11 - 3). They ranged from 11 - 2 to 17 - 9. The procedure was identical to the one used in Experiment 1.

RESULTS

An amount of 5.78% of all trials was spoiled due to failures of the sound-activated relay. As these invalid trials returned at the end of the block, the amount of trials spoiled due to failures of the sound-activated relay was reduced to 1.08%. Further, all incorrect trials (5.46%), all choice trials on which the ‘other’ strategy was chosen (0.18%), and all no-choice trials on which participants failed to use the forced strategy (11.58%) were deleted. Three univariate analyses of variance (ANOVAs) with loaded working-memory component (passive phonological, active phonological with 3 letters, active phonological with 5 letters, executive) as between-subjects effect were conducted to test possible differences across the four groups. A first ANOVA, conducted on the scores on the French Kit, showed that the four groups did not differ in mathematical skill, $F < 1$ (means of 32.3, 29.7, 30.9, and 25.6, respectively). A second ANOVA, conducted on the scores of the calculator-use questionnaire, showed that the four groups did not differ in their reported calculator use, $F(3,36) = 1.14$ (means of 2.8, 3.6, 3.5, and

3.5). The last ANOVA, conducted on the number of weekly arithmetic lessons in the last year of secondary school, showed no differences across groups either, $F(3,36) = 1.49$ (means of 4.1, 5.0, 5.0, and 5.4).

Strategy efficiency. A 4 x 2 x 3 ANOVA was conducted on no-choice RTs with working-memory component (passive phonological, active phonological with 3 letters, active phonological with 5 letters, executive) as between-subjects effect and load (no load vs. load) and strategy (retrieval, transformation, counting) as within-subjects effects (see Table 4). All main effects were significant. RTs were higher under load than under no-load, $F(1,36) = 23.57$, $MSe = 187564$. The main effect of working-memory component, $F(3,36) = 5.96$, $MSe = 1232430$, indicated that RTs were higher under executive load than under phonological load (all phonological tasks taken together), $F(1,36) = 16.14$, $MSe = 187565$, whereas there was no difference between the three sorts of phonological load (all Fs < 1). As a matter of fact, these effects of working-memory component were restricted to the load sessions, i.e., the load x working-memory component interaction was significant, $F(3,36) = 5.53$. More specifically, the effect of load was significant for the executive working-memory component, $F(1,36) = 34.89$, but not for any of the phonological working-memory components (highest $F = 2.78$). The main effect of strategy, finally, $F(2,35) = 181.76$, $MSe = 915406$, indicated that RTs differed significantly across strategies, with retrieval RTs being smaller than transformation RTs, $F(1,36) = 209.83$, $MSe = 96614$, which were in their turn smaller than counting RTs, $F(1,36) = 156.38$, $MSe = 1283826$.

Load and strategy interacted, $F(2,35) = 12.74$, $MSe = 1407550$. Although the effect of load was significant for each single strategy, it was larger on transformation than on retrieval, $F(1,36) = 15.35$ and larger on counting than on transformation, $F(1,36) = 5.78$.

The three-way interaction between load, working-memory component, and strategy tended to be significant, $F(6,72) = 1.98$ ($p = .08$).

Separate ANOVAs for each single strategy tested the effects of the different working-memory loads. The effect of an executive working-memory load was significant for all three strategies, whereas the active phonological load with 5 letters did affect transformation RTs only. The active phonological task with 3 letters had no effect at all, but counting RTs were significantly affected by the passive phonological task. The ANOVA on all strategies showed that the effect of an executive load was higher in counting than in transformation and higher in transformation than in retrieval. Also important, transformation RTs were still more affected by the executive load than by the active phonological load (as measured with the 5-letter task).

Table 4 (continued on next page)
No-choice RTs (in msec) as a function of load, working-memory component[a], and strategy. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Retrieval	No load	915 (75)	1002 (75)	1092 (75)	1151 (75)	1040 (37)
	Load	915 (60)	1071 (60)	1001 (60)	1362 (60)	1087 (30)
	Mean	915 (63)	1036 (63)	1047 (63)	1256 (63)	1064 (32)

[a] PHON = phonological, CE = central executive.

Table 4
No-choice RTs (in msec) as a function of load, working-memory component[a], and strategy. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Transformation	No load	1590 (106)	1593 (106)	1755 (106)	1757 (106)	1674 (53)
	Load	1658 (121)	1680 (121)	1895 (121)	2275 (121)	1877 (60)
	Mean	1624 (106)	1637 (106)	1825 (106)	2016 (106)	1775 (53)
Counting	No load	3558 (407)	3140 (407)	3781 (407)	4455 (407)	3734 (203)
	Load	4049 (394)	3453 (394)	3981 (394)	5708 (394)	4298 (197)
	Mean	3804 (371)	3297 (371)	3881 (371)	2082 (371)	4016 (186)
All strategies	No load	2021 (159)	1912 (159)	2209 (159)	2454 (159)	2149 (79)
	Load	2207 (149)	2068 (149)	2292 (149)	3115 (149)	2421 (74)
	Mean	2114 (143)	1990 (143)	2251 (143)	2785 (143)	2285 (72)

[a] PHON = phonological, CE = central executive.

Strategy selection. In order to investigate effects on strategy selection, a 4 x 2 x 3 ANOVA was conducted on percentage strategy use (in the choice condition), with working-memory component as between-subjects effect and load and strategy as within-subjects effects. As in Experiment 1, all three strategies were used spontaneously by the participants (see Table 5). The main effect of strategy, $F(2,35) = 218.59$, $MSe = 1288$, indicated that retrieval (60%) was used more often than transformation (34%), $F(1,36) = 11.64$, $MSe = 2365$, which was in its turn used more frequently than counting (6%), $F(1,36) = 42.58$, $MSe = 758$. The main effects of load and working-memory component did not reach significance, and neither did any interaction (highest $F = 1.3$).

Table 5
Strategy use in the choice condition (in %) as a function of load and working-memory component[a] for all three strategies. Standard errors are shown between brackets.

		PHON passive	PHON active 3 letters	PHON active 5 letters	CE	Mean
Retrieval	No load	62 (8)	52 (8)	70 (8)	56 (8)	60 (4)
	Load	68 (8)	48 (8)	74 (8)	53 (8)	61 (4)
Transformation	No load	30 (8)	45 (8)	23 (8)	39 (8)	34 (4)
	Load	27 (8)	49 (8)	16 (8)	42 (8)	34 (4)
Counting	No load	8 (3)	2 (3)	8 (3)	5 (3)	6 (1)
	Load	4 (3)	3 (3)	9 (3)	5 (3)	5 (2)

[a] PHON = phonological, CE = central executive.

Secondary task performance. Secondary task performance can be found in Table 3. Performance on the CRT task was significantly faster and more accurate when executed alone than when executed simultaneously with the primary task (choice and no-choice conditions taken together), $F(1,9) = 13.84$, $MSe = 5460$ and $F(1,9) = 241.47$, $MSe = 88$, respectively. There was no difference in CRT speed across the choice and no-choice conditions, but CRT accuracy was lower in the no-choice/count condition than in no-choice/retrieval and no-choice/transform conditions, $F(1,9) = 4.70$, $MSe = 60$ and $F(1,9) = 5.10$, $MSe = 60$, respectively.

Performance on the active phonological tasks with 3 and 5 letters was more accurate when executed alone than when executed simultaneously with the primary task (choice and no-choice conditions taken together), $F(1,9) = 26.51$, $MSe = 188$ and $F(1,9) = 8.48$, $MSe = 162$, respectively. Performance on the 3-letter task was also more accurate in the no-choice/transform condition than in choice, no-choice/retrieval, and no-choice/count conditions, $F(1,9) = 14.14$, $MSe = 130$, $F(1,9) = 5.31$, $MSe = 126$, $F(1,9) = 8.64$, $MSe = 80$, respectively. There were no differences in performance on the 5-letter task across the choice condition and the three no-choice conditions.

SUMMARY

The results obtained in Experiment 1 were generalized to subtraction problems, since participants' performances were shown to be affected by an executive working-memory load. An active phonological working-memory load (as measured by the 5-letter task) affected performance as well, albeit only when transformation strategies were used. The present data also showed that a passive phonological load affected counting RTs in subtraction problems. Furthermore, procedural strategies were shown to be more heavily affected by executive and phonological working-memory loads than retrieval strategies were. Concerning strategy selection, finally, present results showed that retrieval was the most frequently used strategy, followed by transformation, whereas counting was used rather rarely. No effect of working-memory load on strategy selection was observed.

GENERAL DISCUSSION

All strategies (i.e., retrieval, transformation, and counting) were performed less efficiently under an executive working-memory load, in both addition and subtraction problems. However, the degree to which the different strategies were affected differed. As the effects of an executive working-memory load were significantly smaller on retrieval RTs than on procedural RTs, we might conclude that direct memory retrieval required rather few executive working-memory resources, whereas the procedural strategies counting and transformation required substantial amounts of executive working-memory resources. For addition problems, an active phonological load (the 5-letter task) affected both retrieval and procedural RTs, but to a larger extent the latter than the former. For subtraction problems, in contrast, an active phonological load affected transformation RTs only. A passive phonological load finally, only affected RTs when counting was used to solve subtraction problems.

To summarize, executive working-memory resources played a role in retrieval and procedural efficiency. Active phonological working-memory resources played a role in procedural efficiency under some conditions but were unrelated to retrieval efficiency. However, these conclusions should be treated with caution. One has to keep in mind that only one measure of arithmetic strategy efficiency was examined, namely speed. Accuracy was not included because error rates were very low. The results on strategy efficiency obtained in the current study thus concern strategy speed and not strategy accuracy. More specifically, if participant's performance was slowed down by a specific working-memory load, we can conclude that this specific working-memory component was needed to solve simple-arithmetic problems *quickly*. Whether or not this working-memory component is needed to solve simple-arithmetic problems *accurately* remains an open question that future research might resolve. It is, however, difficult to use adults' simple-arithmetic accuracy data, as error rates are usually very low. In the following, we address the question which functions the central executive and the phonological loop might fulfill in simple arithmetic.

THE ROLE OF THE CENTRAL EXECUTIVE

An executive working-memory load affected both retrieval and procedural RTs. In *procedural* strategies, executive working-memory resources are needed to select and implement the appropriate heuristics when the solution is not directly available through retrieval, and to perform the calculations required for mental arithmetic (e.g., Ashcraft, 1995; Imbo, Vandierendonck, & De Rammelaere, in press e; Imbo, Vandierendonck, & Vergauwe, in press g; Logie, Gilhooly, & Wynn, 1994). The manipulation of interim results during calculation would also be controlled by the central executive (Fürst & Hitch, 2000). The fact that the central executive is needed to monitor the number just said and the next count (e.g., Case, 1985; Hecht, 2002; Logie & Baddeley, 1987) may explain why counting needed even more executive resources than transformation did. Keeping track of counted and to-be-counted items and keeping track of one's progress in a counting sequence indeed places demands on the central executive (Ashcraft, 1995).

The significant effect of an executive working-memory load on *retrieval* RTs implies a possible role for the central executive in memory retrieval. This result is in agreement with results obtained recently by Barrouillet, Bernardin, and Camos (2004), who observed that cognitive resources are needed to perform even the simplest retrievals of over-learned knowledge from long-term memory. However, whether or not the central executive is needed in direct memory retrieval remains a debated topic. Retrieving an answer form long-term memory is composed of two processes.

First, several candidate answers – which are represented in an interrelated network of associative links in long-term memory – are automatically activated (e.g., Ashcraft, 1992; Campbell, 1995). Second, one of these answers should be selected as the correct one. One may question whether these processes need executive working-memory resources to be executed.

It has been suggested that the interaction between working memory and long-term memory is one of the functions of the central executive (e.g., Baddeley, 1996). If the central executive indeed plays a crucial role in the activation of information in long-term memory, then it is very likely that people with reduced working-memory space (either by a low working-memory capacity, by a state in which working memory is loaded, or by a physiological cause) will experience difficulties in fact retrieval (e.g., Ashcraft, 1995; Conway & Engle, 1994; Kaufmann, 2002). Consequently, insufficient activation of the correct problem-answer association may slow down retrieval processes (Ashcraft, 1995).

The second process, choosing one answer as the correct one, may also load executive working-memory resources. Deschuyteneer and colleagues (2005a, 2005b, 2006, 2007), for example, showed that the executive working-memory functions "response selection" and "inhibition" are important constituents to solve simple-arithmetic problems. Because retrieving a correct answer to an arithmetic problem involves selecting this answer and inhibiting several similar answers (or 'neighbors'), executive working-memory resources are needed to resolve this competition between the correct answer and its neighbors, and to select the correct response.

In spite of these explanations, the question whether the elementary process of fact retrieval does rely on the central executive is still a debated topic. Some authors do believe that executive working-memory functions are needed (e.g., Ashcraft, 1995; Baroody, 1994; De Rammelaere & Vandierendonck, 2001; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2000, 2002), whereas others don't. Hecht (2002), for example, maintains that retrieval does occur automatically (i.e., without relying on any working-memory component), even though he observed slower retrieval times when working memory was executively loaded than when it was not loaded. Such effects could indeed also be due to general processes such as comparison and decision effects. In their review, DeStefano and LeFevre (2004) also defend that the use of the central executive in retrieval is tied to general attentional requirements of the task. Although present results do not resolve this ongoing discussion, they provide some guidelines for further research. A more detailed analysis of different executive working-memory functions and different arithmetic strategies might be an interesting line for future studies. A combination of the choice/no-choice method and various secondary tasks loading different executive working-memory functions would be an excellent methodology for such a study.

THE ROLE OF THE PHONOLOGICAL LOOP

The phonological task with 5 letters, which loaded the active phonological rehearsal process, affected the retrieval strategy in the addition experiment but not in the subtraction experiment. The phonological task with 5 letters also affected procedural strategies, although the effect of this task did not reach significance when counting was used to solve subtraction problems. The active rehearsal process may indeed play several roles in arithmetic procedures, such as keeping track of running totals and temporarily storing intermediate or partial results (e.g., Ashcraft, 1995; Heathcote, 1994; Imbo et al., in press e,g; Logie & Baddeley, 1987; Logie et al., 1994; Seitz & Schumann-Hengsteler, 2002). In the addition experiment, the role of the active phonological rehearsal process was even more important in counting than in transformation, which is in agreement with previous studies in which counting processes were investigated in a more direct manner (e.g., counting of dots). Logie and Baddeley (1987), for example, observed that counting processes were significantly disrupted by phonological working-memory loads. Camos and Barrouillet (2004) observed longer counting times under a phonological working-memory load (maintaining 5 items in memory) as well. Logie and Baddeley (1987) further state that mental counting involves ‘subvocal articulation of numbers in the counting sequence’, whereas Ashcraft (1995) concludes that the phonological loop would be especially involved in counting, given the phonological basis of the one-by-one incrementing process. It is not clear why no significant effects of the active phonological task with 3 letters were observed. The most plausible explanation is that retaining 3 letters in memory was not demanding enough to affect the arithmetic performance. Otherwise stated, although retaining 3 letters in memory must have loaded the active subvocal rehearsal process, there must have been enough space left to retain digits in memory as well (see also Baddeley & Hitch, 1974).

The passive phonological store (which was loaded by irrelevant speech) was needed when counting was used to solve subtraction problems. This observation is consistent with the assertion that the passive phonological store is used to continually register the running total obtained by the subsequent counting steps. This continued registration of the running total was only needed in counting down processes (9…8…7) but not in counting up processes (7…8…9), however. This dissociation might be caused by the fact that counting up is over-learned and occurs rather automatically, as opposed to counting down. Indeed, counting down is contra-intuitive and people may need to register each count to preclude themselves from forgetting which is the current digit in the counting sequence. Consequently, in the passive phonological store the irrelevant speech (which gains direct access to the store) might have affected the subvocal articulation of each digit. A significant role of the passive phonological store in counting has been observed by Logie and Baddeley (1987) as well.

The significant role of the active rehearsal process (as measured by the active phonological task with 5 letters) in *retrieval* strategies was rather unexpected, because there is no specific reason to assume that phonological working-memory resources are needed in direct memory retrieval. We propose that this result might have been caused by a methodological artifact of our study. More specifically, because maintaining 5 letters in memory is quite hard, executive functions might have come into play as well, e.g., to compare old with new information and to decide whether the letters are similar or not. The active phonological task with 5 letters may thus not have been a purely phonological task. The lower retrieval efficiency under the active phonological task with 5 letters can then be explained by a feature of the design, and not by a significant role of the rehearsal process in retrieval strategies.

Future research is needed to investigate which secondary tasks are most suited to study the role of phonological working-memory resources in simple-arithmetic performance. Researchers should search for a task in which phonological items should be retained in memory for a certain period (i.e., more than 2 seconds; otherwise passive phonological resources may fulfill the task). The amount of items to be held in memory should not be exaggerated either, since overloading the rehearsal process may call executive processes into play. The present results suggest that a memory load of 3 letters is somewhat too small whereas a memory load of 5 letters would already overload the active rehearsal process.

CONCLUSION

The present results showed that retrieval, transformation, and counting strategies are slowed down by an executive working-memory load. Efficient strategic performance might thus rely on executive resources. Procedural strategies were also slowed down by an active phonological load, whereas only counting efficiency was affected by a passive phonological load. Finally, strategy selection was not affected by any working-memory load. Future research might elaborate on these results and (a) investigate *which* executive functions (e.g., inhibition, memory updating, response selection, …) are needed in efficient strategic performance, (b) investigate whether or not working memory is needed in other aspects of strategic performance (e.g., strategy accuracy, strategy adaptivity, …), and (c) test whether the results obtained in the current study generalize to other operations and/or more complex-arithmetic problems.

CHAPTER 5
DO MULTIPLICATION AND DIVISION STRATEGIES RELY ON EXECUTIVE AND PHONOLOGICAL WORKING-MEMORY RESOURCES?

Memory & Cognition (in press)[1,2]

The role of executive and phonological working-memory resources in simple arithmetic was investigated in two experiments. Participants had to solve simple multiplication problems (e.g., 4 x 8; Experiment 1) or simple division problems (e.g., 42 : 7; Experiment 2) under no-load, phonological-load, and executive-load conditions. The choice/no-choice method was used to investigate strategy execution and strategy selection independently. Results on strategy execution showed that executive working-memory resources were involved in direct memory retrieval of both multiplication and division facts. Executive working-memory resources were also needed to execute nonretrieval strategies. Phonological working-memory resources, in contrast, tended to be involved in nonretrieval strategies only. Results on strategy selection showed no effects of working-memory load. Finally, correlation analyses showed that both strategy execution and strategy selection correlated with individual-difference variables such as gender, math anxiety, associative strength, calculator use, arithmetic skill, and math experience.

---

[1] This paper was co-authored by André Vandierendonck.

[2] Thanks are extended to Mark Ashcraft, Sian Beilock, and one anonymous reviewer for their helpful comments on previous drafts of this article.

INTRODUCTION

Working memory is a system devoted to short-term storage and processing, and is used in various cognitive tasks such as reading, reasoning, and mental arithmetic. The last decennia, research into the role of working memory in mental arithmetic has flourished (for review, see DeStefano & LeFevre, 2004) and showed that solving both simple-arithmetic problems (e.g., 8 + 5, 3 x 9) and complex-arithmetic problems (e.g., 23 + 98, 12 x 35) relies on working-memory resources. The present study further investigates the role of working memory in simple-arithmetic strategies, based on the multi-component working-memory model of Baddeley and Hitch (1974). In this model there is an attentional system (the central executive) that supervises a phonological subsystem and a visuo-spatial subsystem, which guarantee short-term maintenance of phonological and visuo-spatial information, respectively.

The role of executive working-memory resources in simple arithmetic has been shown extensively (e.g., Ashcraft, 1995; De Rammelaere, Stuyven, & Vandierendonck, 1999, 2001; De Rammelaere & Vandierendonck, 2001; Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Coeman, 2007; Deschuyteneer, Vandierendonck, & Muyllaert, 2006; Hecht, 2002; Lemaire, Abdi, & Fayol, 1996; Seitz & Schumann-Hengsteler, 2000, 2002). The role of phonological working-memory resources in simple arithmetic is less clear. In some studies an effect of phonological load on simple-arithmetic problem solving was observed (e.g., Lee & Kang, 2002; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2002) whereas in other studies it was not (e.g., De Rammelaere et al., 1999, 2001; Seitz & Schumann-Hengsteler, 2000). The role of the visuo-spatial sketchpad in simple arithmetic has only scarcely been investigated (but see Lee & Kang, 2002; Seitz & Schumann-Hengsteler, 2000) and is equivocal.

A drawback of all the studies mentioned above, however, is that none of them made a separation between retrieval and nonretrieval trials. Yet, it has been shown that adults use several strategies to solve even the simplest arithmetic problems (e.g., Hecht, 1999; LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). More specifically, although direct memory retrieval (i.e., 'knowing' that 3 x 4 = 12) is the most frequently used strategy, nonretrieval strategies (or procedural strategies) are used as well. Examples of such procedural strategies are transformation (e.g., 9 x 6 = (10 x 6) – 6 = 60 – 6 = 54) and counting (e.g., 4 x 7 = 7… 14… 21… 28). Based on the studies mentioned above, it is impossible to know *in which* simple-arithmetic strategies executive and phonological working-memory resources are needed.

The role of executive and phonological working-memory across different simple-arithmetic strategies started to be investigated only very recently. Hecht (2002) conducted the first study on this topic. In his study, simple addition equations (e.g., 4 + 3 = 8, true/false?) had to be verified under no load, phonological load, and executive load. After each trial, participants had to report which strategy they had used. Results showed that all strategies (i.e., retrieval, transformation, and counting) were slowed down under executive working-memory loads, whereas only the counting strategy was slowed down under phonological working-memory loads. Based on regression analyses however, Hecht concluded that retrieval does not rely on the central executive, whereas the counting strategy would rely on both executive and phonological working-memory resources.

Seyler, Kirk, and Ashcraft (2003) also studied the role of working memory in simple-arithmetic strategies. In a first experiment, simple subtraction problems had to be solved while a 2-, 4-, or 6-letter string had to be remembered. Results showed that solving subtraction problems with minuends of 11 or greater (e.g., 11 – 5) relied more heavily on working memory than problems with minuends smaller than 11 (e.g., 8 – 5). In another experiment, using strategy reports, Seyler et al. (2003) showed that subtraction problems with minuends of 11 or greater were more frequently solved with procedural strategies than problems with minuends smaller than 11. It was concluded that working memory is more heavily involved when simple subtraction problems are solved via procedural strategies.

A drawback of both previous studies is that neither Hecht (2002) nor Seyler et al. (2003) controlled for strategy selection effects, since participants were always free to choose any strategy they wanted. Consequently, nonretrieval strategies will have been executed more frequently on large problems while retrieval will have been executed more frequently on small problems. Such strategy selection effects might have influenced strategy efficiency data and all resulting conclusions. In order to exclude such biasing effects of strategy selection on strategy efficiency, the choice/no-choice method (Siegler & Lemaire, 1997) should be used. Using the choice/no-choice method in combination with selective working-memory loads provides unbiased data about the role of working memory in strategy selection and strategy efficiency.

The combination of the choice/no-choice method and selective working-memory loads has first been used by Imbo and Vandierendonck (in press b). They investigated the role of executive and phonological working-memory resources in simple-arithmetic strategies. In their study, simple addition and subtraction problems had to be solved under no-load, passive-phonological load, active-phonological load, or central-executive load conditions. Results showed that retrieval of addition and subtraction facts relied on executive working-memory resources. Solving addition or subtraction problems by means of a nonretrieval strategy required both executive and active-phonological working-memory resources. The passive phonological store was only involved when counting was used to solve subtraction problems. Obviously, the role of executive and phonological working-memory resources was significantly larger in nonretrieval strategies (i.e., transformation and counting) than in direct memory retrieval.

To summarize, the three studies described above showed that the role of working memory differs across strategies. Whether or not the central executive is needed in retrieval remains a debated topic: Hecht (2002) does not believe that this working-memory component is needed in retrieval whereas Imbo and Vandierendonck (in press b) presented evidence that retrieval requires executive working-memory resources. Nevertheless, all three studies seem to agree that phonological working-memory resources are needed when nonretrieval strategies are used to solve simple addition and/or subtraction problems.

Our knowledge about the role of working memory in addition and subtraction strategies may be scarce; the knowledge about the role of working memory in multiplication and division strategies is non-existent. Despite that fact that solving simple multiplication and division problems requires working-memory resources (De Rammelaere & Vandierendonck, 2001; Deschuyteneer & Vandierendonck, 2005b; Deschuyteneer et al., 2006, 2007; Lee & Kang, 2002; Seitz & Schumann-Hengsteler, 2000, 2002), up until now, no study investigated the role of working memory across the different multiplication and division strategies.

As multiplication and division can certainly not be seen as the counterparts of addition and subtraction, studying the role of working memory in multiplication and division strategies is much more than merely an extension of previous research. Indeed, there exist many differences across operations; and especially between addition and subtraction, on the one hand, and multiplication and division, on the other. Differences across arithmetic operations start from childhood on and continue to exist in adulthood. First, addition and subtraction problem-solving procedures are taught before multiplication and division problem-solving procedures. Furthermore, the acquisition of addition and subtraction is mainly based on counting procedures, whereas the acquisition of multiplication and division is based on the memorization of problem-answer pairs. In adults, the highest percentages retrieval use are observed in multiplication (98%) whereas the lowest percentages retrieval use are observed in division (69%), with addition and subtraction in between (88% and 72%, respectively; Campbell & Xue, 2001). Adults' strategy efficiencies differ also greatly across operations, with multiplication RTs (930 msec) being much faster than division RTs (1086 msec, Campbell & Xue, 2001).

These results seem to suggest that access to long-term memory and selecting the correct response is very difficult for division but rather easy for multiplication. As getting access to long-term memory and selecting the correct response are processes requiring *executive* working-memory resources, one might rather be sure that an executive load will affect division efficiency, but it might be questioned whether an executive load will affect the over-learned retrieval of multiplication facts. It might further be expected that *phonological* working-memory loads will affect nonretrieval strategy efficiencies but not retrieval strategy efficiencies. Indeed, when nonretrieval strategies are used, intermediate values have to be kept temporary in working memory, a function accomplished by the phonological working-memory component (Ashcraft, 1995). Effects of phonological working-memory loads on nonretrieval strategies have been observed in addition and subtraction, but as several authors (e.g., Campbell, 1994; Dehaene, 1997) suppose that multiplication is more heavily based on auditory-verbal number codes than other operations are, effects of phonological working-memory loads may be more heavily apparent in the present study.

In order to investigate the role of executive and phonological working-memory resources[3] in multiplication and division strategies, the present study used two frequently used and approved methods: the selective interference paradigm and the choice/no-choice method. The selective interference paradigm is the methodological approach most frequently

[3] Given the poorer elaboration of the role of the visuo-spatial sketchpad in simple arithmetic (on theoretical, methodological, and empirical level), this working-memory component was not included in the present study.

chosen for studying the role of different working-memory resources in mental arithmetic. It entails using both a single-task condition in which the primary task (mental arithmetic) is executed without any working-memory load and a dual-task condition in which the primary task is combined with a secondary task loading a specific working-memory component. If both primary and secondary task demand the same working-memory resources, performance decrements should be observed in either task. In the present study, three secondary tasks were used to load three specific working-memory components, more specifically the passive-phonological component (the phonological store), the active-phonological component (the subvocal rehearsal process), and the central executive.

The choice/no-choice method (designed by Siegler & Lemaire, 1997) is used to collect data on strategy selection (which strategies are chosen?) and strategy efficiency (are strategies executed efficiently?) independently. In this method, each participant is tested under two types of conditions: a choice condition in which participants are free to choose any strategy they want and no-choice conditions in which participants are required to solve all the problems with one particular strategy. There are as many no-choice conditions as there are strategies available in the choice condition. Data obtained in no-choice conditions provide information about strategy efficiency, whereas data gathered in the choice condition provide information about strategy selection.

Besides investigating the role of working memory in multiplication and division strategies, the present study also wanted to test whether simple-arithmetic strategies are influenced by factors *not* imposed by the experimenter. To this end, several individual-difference measures were obtained for each participant, namely arithmetic skill, math experience, gender, calculator use, math anxiety, and associative strength. Effects of *arithmetic skill* have already been reported (e.g., Campbell & Xue, 2001; Gilles, Masse, & Lemaire, 2001; Kirk and Ashcraft, 2001; LeFevre &

Bisanz, 1986; LeFevre et al., 1996a, 1996b). Generally, strategy use is more efficient (i.e., faster) in high-skill participants than in low-skill participants.

Effects of *math experience*, in contrast, have been reported only rarely. Roussel, Fayol, and Barrouillet (2002) observed that experienced participants (primary school teachers) performed slower on arithmetic tasks than did inexperienced participants (undergraduate psychology students). In contrast, experienced and inexperienced participants did not differ in their strategy choices. In one of our own studies, arithmetic experience (based on the participants' secondary school curricula) was found to predict both strategy selection and strategy efficiency, albeit only for multiplication problems and not for addition problems (Imbo, Vandierendonck, & Rosseel, in press f).

Gender effects have been investigated in children rather than in adults. Several child studies showed more frequent and more efficient retrieval use in boys than in girls (e.g., Carr & Jessup, 1997; Carr, Jessup, & Fuller, 1999; Royer, Tronsky, Chan, Jackson, & Marchant, 1999a). Whether or not these differences exist in adulthood, is a debated topic. Because some recent studies (e.g., Geary, Saults, Liu & Hoard, 2000b; Imbo et al., in press f) observed significant gender differences in adults' arithmetic processing, with males outperforming females, gender was included in the present study as well.

Only two studies investigated the possible effects of *calculator use*, one observing no effects (Campbell & Xue, 2001) and one observing effects of calculator use on strategy efficiency (Imbo et al., in press f); participants who reported highly frequent calculator use were remarkably slower in both retrieval efficiency and procedural efficiency. The present study elaborated on this issue and included a short questionnaire about calculator use.

Concerning *math anxiety*, it was expected that high-anxious participants would perform worse on the simple-arithmetic tasks than low-anxious participants. Effects of math anxiety have previously been shown in complex-arithmetic tasks (e.g., Ashcraft & Kirk, 2001) but not yet in simple-arithmetic tasks.

Finally, the *associative strength* variable is an estimate of how strong the participant's problem-answer associations are in long-term memory, and is operationalized as the participant's percentage retrieval use in choice conditions. It was hypothesized that participants with stronger problem-answer associations would be faster in retrieving arithmetic facts from long-term memory.

EXPERIMENT 1: MULTIPLICATION

METHOD

Participants. Sixty subjects participated in the present experiment (15 men and 45 women). Their mean age was 21 years and 0 months. Half of them were first-year psychology students at Ghent University who participated for course requirements and credits. The other half was paid €10 for participation.

Procedure. Each participant was tested individually in a quiet room for approximately 1 hour. The experiment started with short questions about the age of the participant, his/her math experience (i.e., the number of mathematics lessons per week during the last year of secondary school), calculator use (on a rating scale from 1 "never" to 5 "always"), and math anxiety (on a rating scale from 1 "low" to 5 "high"[4]). All participants solved the simple-arithmetic problems in four conditions: first the choice condition

[4] The correlation between rating math anxiety on a scale from 1 to 5 and rating math anxiety with the Mathematics Anxiety Rating Scale (MARS, Richardson & Suinn, 1972) ranges from .45 to .85 (Mark Ashcraft, personal communication).

(in order to exclude influence of no-choice conditions on the choice condition), and then three no-choice conditions, the order of which was randomized across participants. In the choice condition, 6 practice problems and 42 experimental problems were presented. After the choice condition, participants needed no more practice; the no-choice conditions thus immediately started with 42 experimental problems. Each condition was further divided in two blocks: one in which no working-memory component was loaded, and one in which one working-memory component was loaded. The working-memory load differed across participants: for 20 participants the central executive was loaded, for 20 participants the active phonological rehearsal process was loaded, and for 20 participants the passive phonological store was loaded. For half of the participants, each condition started with the no-load block and was followed by the working-memory load block; the order was reversed for the other half of the participants.

Simple-arithmetic task. The multiplication problems presented in the simple-arithmetic task consisted of two one-digit numbers (e.g., 6 x 7). Problems involving 0, 1, or 2 as an operand (e.g., 5 x 0, 1 x 4, 2 x 3) and tie problems (e.g., 3 x 3) were excluded. Since commuted pairs (e.g., 9 x 4 and 4 x 9) were considered as two different problems, this resulted in 42 multiplication problems (ranging from 3 x 4 to 9 x 8). Small problems were defined as problems with a correct product smaller than 25 whereas large problems were defined as problems with a correct product larger than 25 (Campbell, 1997; Campbell & Xue, 2001). A trial started with a fixation point for 500 msec. Then the multiplication problem was presented horizontally in the center of the screen, with the operation sign at the fixation point. The problem remained on screen until the subject responded. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, participants wore a microphone that was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 msec. On each trial, feedback was presented to the participants, a green ‘Correct’ when their answer was correct, and a red ‘Incorrect’ when it was not.

Immediately after solving each problem, participants in the choice condition were presented four strategies on the screen (see e.g. Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler et al., 2003): Retrieval, Counting, Transformation, and Other. These four choices had been extensively explained by the experimenter: *(1) Retrieval: You solve the problem by remembering or knowing the answer directly from memory. It means that you know the answer without any additional processing. For example: you know that 5 x 6 = 30 because 30 "pops into your head". (2) Counting: You solve the problem by counting a certain number of times to get the answer. You recite the tables of multiplication. For example: 4 x 7 = 7… 14… 21… 28 or 5 x 3 = 5… 10… 15. (3) Transformation: You solve the problem by referring to related operations or by deriving the answer from known facts. You change the presented problem to take advantage of a known arithmetical fact. For example: 9 x 8 = (10 x 8) – 8 = 80 – 8 = 72 or 6 x 7 = (6 x 6) + 6 = 36 + 6 = 42. (4) Other: You solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem. For example: guessing.* After each problem, participants were asked to report verbally which of these strategies they had used. The experimenter also emphasized that the presented strategies were not meant to encourage use of a particular strategy. If the participant felt like using only one of the presented strategies, he/she was completely free to do so; when the participant acknowledged generally using a mix of strategies; he/she was as free to do so.

In the no-choice conditions, participants were asked to use one particular strategy to solve all problems. In no-choice/retrieval, they were required to retrieve the answer. More specifically, participants were asked to pronounce the answer that first popped into their head. In no-choice/transformation, participants were required to transform the problem by making an intermediate step. The experimenter proposed several intermediate steps, and all participants recognized using at least a few of them. Examples were (a) going via 10, e.g., 9 x 6 = (10 x 6) - 6 = 60 - 6 = 54 and 5 x 7 = (10 x 7) : 2 = 70 : 2 = 35, (b) using the double, e.g., 4 x 6 = 2 x 2 x 6 = 2 x 12, and (c) using ties, e.g., 7 x 8 = (7 x 7) + 7 = 49 + 7 = 56. However, if participants normally used any transformation step not proposed by the experimenter, they were free to do so. In no-choice/counting, participants had to say (sub-vocally) the tables of multiplication until they reached the correct total (e.g., 4 x 7 = 7… 14… 21… 28). After having solved the problem, participants also had to answer with ‘yes’ or ‘no’ whether they had succeeded in using the required strategy. This enabled us to exclude non-compliant trials from analyses.

In choice and no-choice conditions, the answer of the participant, the strategy information, and the validity of the trial were recorded online by the experimenter. All invalid trials (e.g., failures of the voice-activated relay) were discarded and returned at the end of the block, which minimized data loss due to unwanted failures.

Executive secondary task. A continuous choice reaction time task (CRT task) was used to load the executive working-memory component (Szmalec, Vandierendonck, & Kemps, 2005). Stimuli of this task consisted of low tones (262 Hz) and high tones (524 Hz) that were sequentially presented with an interval of 900 or 1500 msec. Participants had to press the 4 on the numerical keyboard when they heard a high tone and the 1 when a low tone was presented. The duration of each tone was 200 msec. The tones were presented continuously during the simple-arithmetic task. The CRT task was also performed alone (i.e., without the concurrent solving of arithmetic problems). In this single-task condition, the multiplication problems with their correct answer were presented, which the participants had to read off the screen. Doing so, the procedure and vocalization of the task remained very similar to the procedure and vocalization in the dual-task condition. Differences in the secondary-task performance could thus only be due to the mental-arithmetic process itself.

Active phonological secondary task. In this task, letter strings of 3 consonants (e.g., T K X) were read aloud by the experimenter. Known letter strings (e.g., BMW, LSD) were avoided. The participant had to retain these letters and repeat them after three simple-arithmetic problems. Following the response of the participant, the experimenter presented a new 3-letter string. This task was also tested individually (i.e., without the concurrent solving of arithmetic problems), using the same methodology as in the CRT single-task condition.

Passive phonological secondary task. In this task, irrelevant speech was presented to the participants. This speech consisted of dialogues between several people in the Swedish language, which were taken from a compact disc used in language courses. The Swedish dialogues were presented with an agreeable loudness (i.e., around 70 dB) through the headphones. Because both Swedish and Dutch (i.e., the participants' native language) are German languages, phonemes strongly match between both languages. None of the participants had any knowledge of Swedish.

French Kit. After the simple-arithmetic experiment, all participants completed a pen-and-paper test of complex arithmetic, the French Kit (French, Ekstrom, & Price, 1963). The test consisted of two subtests, one page with complex addition problems and one page with complex subtraction and multiplication problems. Participants were given 2 minutes per page, and were instructed to solve the problems as fast and accurately as possible. The correct answers on both subtests were summed to yield a total score of arithmetic skill.

RESULTS

Of all trials 6.9% was spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 1.8%. Further, all incorrect trials (4.4%), all choice trials on which participants reported having used a strategy 'Other' (0.1%), and all no-choice trials on which participants failed to use the required strategy (8.8%) were deleted. All data were analyzed on the basis of the multivariate general linear model; and all reported results are considered to be significant if $p < .05$, unless mentioned otherwise.

To test whether the three subject groups (i.e., loaded by the passive phonological task, the active phonological task, or the executive task) differed from each other, analyses of variance (ANOVAs) were conducted on the scores on the French Kit[5] ('arithmetic skill'), the scores of the calculator-use questionnaire, the amount of arithmetic lessons in the last year of secondary school ('math experience'), and the scores of the math anxiety questionnaire. Results showed that the groups did not differ in any of these variables; all Fs < 1.2 and all ps > .30.

Strategy efficiency. Only the RTs uncontaminated by strategy choices (i.e., no-choice RTs) will be considered, since only these RTs provide clear data concerning strategy efficiency. A 3 x 2 x 3 x 2 ANOVA was conducted on correct RTs with working-memory component (passive phonological, active phonological, executive) as between-subjects factor and load (no load vs. load), strategy (retrieval, transformation, counting), and size (small vs. large) as within-subjects factors (see Table 1).

The main effects of load, size, and strategy were significant, $F(1,57) = 10.24$, $MSe = 1326374$, $F(1,57) = 198.87$, $MSe = 1598084$, and $F(2,56) = 110.27$, $MSe = 5221560$, respectively. RTs were longer under load (3061 msec) than under no-load (2786 msec), longer for large problems (3588 msec) than for small problems (2259 msec). RTs were also longer for counting (4759 msec) than for transformation (2992 msec), $F(1,57) = 138.10$, $MSe = 3378924$, and longer for transformation than for retrieval

[5] Both subtests of the French kit correlated significantly with each other ($p < .01$); $r = .675$ in Experiment 1 and $r = .531$ in Experiment 2, indicating high reliability. Correlations are not 100% because both subtests test other operations (addition vs. multiplication-subtraction).

(1020 msec), $F(1,57) = 82.98$, $MSe = 4514306$. The main effect of strategy was modified by a strategy x load interaction and a strategy x size interaction. The strategy x load interaction, $F(2,56) = 5.15$, $MSe = 683977$, indicated that the load effect (i.e. load RTs – no-load RTs) was larger for counting than for retrieval, $F(1,57) = 10.04$, $MSe = 750311$, and larger for counting than for transformation, $F(1,57) = 7.01$, $MSe = 807632$. Load effects did not differ between retrieval and transformation, $F(1,57) < 1$. The strategy x size interaction, $F(2,56) = 69.61$, $MSe = 1705536$, indicated that the problem-size effect (i.e., RTs on large problems – RTs on small problems) was larger in counting than in retrieval, $F(1,59) = 141.63$, $MSe = 2275821$, and larger in counting than in transformation, $F(1,59) = 132.01$, $MSe = 262806$, but as large in retrieval as in transformation, $F(1,59) = 2.13$, $MSe = 212582$ ($p = .15$).

The working-memory component x load interaction did not reach significance, $F(2,57) = 1.91$, $MSe = 1326374$ ($p = .16$). However, as differential load effects were predicted for the different working-memory components, planned comparisons were conducted. These analyses showed that the effect of load (i.e., load RTs – no-load RTs) was significant for the executive component, $F(1,57) = 11.59$, $MSe = 1326374$, but did not reach significance for the active phonological component, $F(1,57) = 1.87$ ($p = .18$) or the passive phonological component, $F(1,57) < 1$. This interpretation was verified by separate ANOVAs that tested the effects of the different working-memory loads for each single strategy. Retrieval RTs were affected by an executive load, $F(1,57) = 35.69$, $MSe = 28055$, but not by an active phonological load, $F(1,57) = 2.38$ ($p = .13$) or a passive phonological load, $F(1,57) < 1$. Transformation RTs tended to be affected by an executive load, $F(1,57) = 2.88$, $MSe = 1054430$ ($p = .09$) but not by an active phonological load, $F(1,57) < 1$ or a passive phonological load, $F(1,57) < 1$. Finally, counting RTs were affected by an executive load, $F(1,57) = 10.16$, $MSe = 1611840$, tended to be affected by an active phonological load, $F(1,57) = 2.75$, $MSe = 1611840$ ($p = .10$), and were not affected by a passive phonological load, $F(1,59) = 1.82$ ($p = .18$). High variance on the counting RTs might have prevented this effect to reach significance.

Table 1
Mean correct RTs (in msec) as a function of load, working-memory component[a], size, and strategy. Standard errors are shown between brackets.

		PHON passive		PHON active		Executive	
		No load	Load	No load	Load	No load	Load
Retrieval	Small	854 (52)	843 (58)	922 (52)	977 (58)	736 (52)	957 (58)
	Large	1129 (80)	1089 (78)	1259 (80)	1319 (78)	964 (80)	1191 (78)
Transformation	Small	2874 (357)	2954 (380)	3280 (357)	3240 (380)	2379 (357)	2761 (380)
	Large	3235 (304)	3126 (334)	3110 (304)	3312 (334)	2616 (304)	3013 (334)
Counting	Small	2881 (269)	3162 (292)	2980 (269)	3342 (292)	2556 (269)	2964 (292)
	Large	6261 (661)	6704 (761)	6284 (661)	6863 (761)	5874 (661)	7275 (761)

[a] PHON = phonological.

To consolidate the results described above, and to investigate the influence of individual differences, correlations[6] were calculated between strategy efficiency (i.e., retrieval RTs, transformation RTs, and counting RTs), strategy selection, working-memory load (i.e., executive, active phonological, and passive phonological), problem size, and individual-difference variables (i.e., math anxiety, arithmetic skill, calculator use, gender, and math experience).

When looking at the correlation measures presented in Table 2, we see that strategies were executed more slowly when problem size was higher and when the central executive was loaded, which confirms the ANOVA results. Moreover, the efficiency of the different strategies correlated with several individual-difference variables. The efficiency of all three strategies was higher in high-skill participants than in low-skill participants. Participants with stronger problem-answer associations were more efficient in retrieval but not in transformation and counting. Retrieval efficiency was higher in infrequent calculator users than in frequent calculator users, and higher in males than in females.

Strategy selection. In order to investigate effects on strategy selection, a 3 x 2 x 2 ANOVA was conducted on percentages use of each single strategy (in the choice condition), with working-memory component (passive phonological, active phonological, executive) as between-subjects factor and load (no load vs. load) and size (small vs. large) as within-subjects factors (see Table 3).

---

[6] Gender was coded as a dummy variable: girls were coded as -1 and boys were coded as 1. Each working-memory load was coded a dummy variable as well. This variable was -1 for no-load conditions and 1 for load conditions.

Table 2
Correlation table for Experiment 1 (multiplication).

	Transform RT	Count RT	Retrieval use[1]	Problem size	Arithm. skill	Calculator use	Math experience	Math anxiety	Gender	Phon. passive	Phon. active	Exec.
Retrieval RT	.424*	.393*	-.370*	.311*	-.415*	.294*	-.006	.009	-.210*	-.021	.045	.193*
Transform RT		.509*	-.113	.539*	-.208*	.093	.002	-.051	-.047	.037	.046	.096
Count RT			-.002	.063	-.284*	.109	.006	-.112	.012	-.012	.045	.080
Retrieval use[1]		.		-.349*	.190*	-.205*	.256*	-.202*	.270*	.016	.007	.048
Arithm. skill						-.440*	.014	.012	.410*	--	--	--
Calculator use							.127	.096	-.332*	--	--	--
Math experience								-.455*	.159	--	--	--
Math anxiety									-.186	--	--	--
Gender										--	--	--

[1] Associative strength is operationalized by the participants' percentage retrieval use
* $p < .0038$ (the Bonferroni-corrected *alpha* level of .05 when correlating 13 variables)
$df = 238$

For retrieval, the main effect of size was significant, $F(1,57) = 71.47$, $MSe = 96$, indicating more frequent retrieval use on small problems (89%) than on large problems (72%). The main effects of load and working-memory component did not reach significance, and neither did any interaction (highest $F = 2.31$). For transformation, the main effect of size was significant as well, $F(1,57) = 50.22$, $MSe = 11395$, indicating more frequent transformation use on large problems (16%) than on small problems (3%). None of the other effects reached significance (highest $F = 1.79$). Finally, counting tended to be used more often on large problems (11%) than on small problems (9%), but this effect did not reach significance, $F(1,57) = 3.13$, $MSe = 403$ ($p = .08$). None of the other effects reached significance (highest $F = 1.18$).

Table 3
Mean percentages strategy use as a function of load, working memory component[a], and size. Standard errors are shown between brackets.

		PHON passive		PHON active		Executive	
		No load	Load	No load	Load	No load	Load
Retrieval	Small	88 (4)	90 (4)	87 (4)	86 (4)	88 (4)	91 (4)
	Large	70 (6)	71 (6)	68 (6)	70 (6)	76 (6)	79 (6)
Transformation	Small	2 (2)	2 (4)	4 (2)	3 (1)	2 (2)	2 (1)
	Large	17 (5)	17 (4)	20 (5)	16 (4)	15 (5)	13 (4)
Counting	Small	9 (3)	8 (4)	10 (3)	11 (4)	9 (3)	7 (4)
	Large	13 (3)	12 (3)	12 (3)	14 (3)	10 (3)	8 (3)

[a] PHON = phonological.

In Table 2, the correlations between retrieval frequency, working-memory load, problem size, and individual differences are presented. Percentage retrieval use correlated with problem size but did not correlate with any of the working-memory loads, which confirms the ANOVA results. Percentage retrieval use correlated with all individual-difference variables, however. More specifically, retrieval was more frequently used by high-skill participants than by low-skill participants, by infrequent calculator users than by frequent retrieval users, by more-experienced participants than by less-experienced participants, by low-anxious participants than by high-anxious participants, and by males than by females.

Secondary task performance. An analysis of variance was conducted on CRT accuracy, CRT speed, and letter-task accuracy (see Table 4) with condition as within-subjects variable (single, choice, no-choice/retrieval, no-choice/transformation, and no-choice/counting). CRT speed tended to differ across conditions, $F(4,16) = 2.56$, $MSe = 3862$, ($p = .08$). Participants were faster to react to the tones in the CRT-only condition (626 msec) than in the other conditions (660 msec), but this difference did not reach significance, $F(1,19) = 2.21$, $MSe = 8516$ ($p = .15$). CRT accuracy differed across conditions as well, $F(4,16) = 6.51$, $MSe = 67$. More specifically, CRT accuracy was significantly higher in the CRT-only condition (87%) than in the other conditions (80%), $F(1,19) = 4.17$, $MSe = 167$. When few executive working-memory resources are left, performance was thus impaired not only on the primary task but also on the secondary task. CRT accuracy was also higher in the no-choice/retrieval condition than in the choice condition, $F(1,19) = 7.31$, $MSe = 32$ and than in the other no-choice conditions, $F(1,19) = 7.04$, $MSe = 40$. Note that the slowest CRT performance was observed in the no-choice/transformation condition, i.e., where the effect of an executive load failed to reach significance ($p = .09$, see above). As such, a trade-off between efficient transformation use and efficient CRT performance may account for the insignificant effect of executive load on transformation RTs.

Performance on the active phonological task (i.e. the letter task) differed across conditions as well, $F(4,16) = 12.56$, $MSe = 166$. Accuracy was significantly higher in the single-task condition (84%) than in dual-task conditions (68%), $F(1,19) = 19.91$, $MSe = 210$.

Table 4
Performance on the secondary tasks in Experiment 1 (multiplication) and Experiment 2 (division). Standard errors are shown between brackets.

Experiment 1	Single	Choice	Retrieval	Transform	Count
CRT accuracy (%)	87 (5)	79 (3)	84 (3)	79 (4)	79 (3)
CRT speed (msec)	626 (26)	656 (17)	646 (20)	672 (18)	666 (16)
Letter-task accuracy (%)	84 (3)	56 (3)	75 (4)	68 (5)	74 (4)
Experiment 2	**Single**	**Choice**	**Retrieval**	**Via multiplication**	
CRT accuracy (%)	88 (2)	73 (3)	75 (3)	75 (3)	
CRT speed (msec)	647 (23)	661 (8)	664 (15)	646 (13)	
Letter-task accuracy (%)	90 (2)	62 (5)	78 (4)	77 (4)	

SUMMARY

Results concerning strategy efficiency showed that the role of the different working-memory resources differed across strategies. Executive working-memory resources were needed in all strategies, whereas phonological working-memory resources were especially needed in the counting strategy. Working-memory load did not have any effect on strategy selection. Both strategy efficiency and strategy selection correlated significantly with several individual-difference variables. The interpretation of the possible roles of these individual differences is postponed to the general discussion.

EXPERIMENT 2: DIVISION

Participants. Sixty subjects (10 men and 50 women) participated in the present experiment. Their mean age was 21 years and 4 months. Half of them were first-year psychology students at Ghent University who participated for course requirements and credits. The other half was paid €10 for participation. None of them had participated in Experiment 1.

Stimuli and Procedure. The 43 division problems were the reverse of the multiplication problems used in Experiment 1. The procedure was identical to the one used in Experiment 1, with one exception. It has been shown that only two strategies are frequently used to solve simple division problems (Campbell & Xue, 2001; LeFevre & Morris, 1999; Robinson, Arbuthnott, & Gibbons, 2002): direct memory retrieval and solving the division problem via the related multiplication problem (e.g., solving 48 : 8 via ? x 8 = 48). Therefore, the choices in the choice condition of this experiment were restricted to three: *(1) Retrieval: You solve the problem by remembering or knowing the answer directly from memory. It means that you know the answer without any additional processing. For example: you know that 30 : 6 = 5 because 5 "pops into your head". (2) Via multiplication: You solve the division problem by using the related multiplication problem. For example: when you have to solve 42 : 6, you think about how many times 6 goes into 42, i.e., 6 x ? = 42. You might also check your answer by doing the multiplication 6 x 7 = ?. (3) Other: You* *solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem. For example: guessing.* Accordingly, there were only two no-choice conditions: no-choice/retrieval, in which participants were asked to retrieve the answer, and no-choice/via-multiplication, in which participants were asked to solve the division problem via the related multiplication problem.

RESULTS

Of all trials, 5.6% were spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 1.5%. Further, all incorrect trials (10.0%), all choice trials on which participants reported having used a strategy 'Other' (0.7%), and all no-choice trials on which participants failed to use the required strategy (6.0%) were deleted. The low percentage of 'Other' strategy use confirms that the two strategies allowed in the choice condition (i.e., direct memory retrieval and the via-multiplication strategy) cover the choice pattern generally used by participants when solving simple division problems. All data were analyzed on the basis of the multivariate general linear model; and all reported results are considered to be significant if $p < .05$, unless mentioned otherwise.

To test whether the three subject groups (i.e., loaded by the passive phonological task, the active phonological task, or the executive task) differed from each other, four analyses of variance (ANOVAs) were conducted. Results showed no group differences in arithmetic skill, calculator use, math experience, or math anxiety; all Fs < 1.1 and all ps > .30.

Strategy efficiency. A 3 x 2 x 2 x 2 ANOVA was conducted on correct no-choice RTs with working-memory component (passive phonological, active phonological, executive) as between-subjects factor and load (no load vs. load), strategy (retrieval vs. via multiplication) and size (small vs. large) as within-subjects factors (see Table 5).

Table 5
Mean correct RTs (in msec) as a function of load, working-memory component[a], size, and strategy. Standard errors are shown between brackets.

		PHON passive		PHON active		Executive	
		No load	Load	No load	Load	No load	Load
Retrieval	Small	745 (59)	725 (75)	917 (59)	908 (75)	906 (59)	1210 (75)
	Large	893 (78)	860 (96)	1159 (78)	1131 (96)	1057 (77)	1402 (96)
Via multiplication	Small	1593 (195)	1696 (193)	1590 (195)	1671 (193)	1410 (195)	1764 (193)
	Large	1996 (281)	2246 (321)	1972 (281)	2107 (321)	1930 (281)	2342 (321)

[a] PHON = phonological.

The main effects of load, strategy, and problem size were significant. RTs were longer under load (1505 msec) than under no-load (1304 msec), $F(1,57) = 29.08$, $MSe = 102768$; retrieving division facts (993 msec) was faster than solving them via multiplication (1860 msec), $F(1,57) = 52.84$, $MSe = 1400216$; and small problems (1261 msec) were solved faster than large problems (1591 msec), $F(1,57) = 59.60$, $MSe = 219528$.

Strategy further interacted with problem size and with load. The strategy x size interaction indicated a larger problem-size effect (i.e., RTs on large problems – RTs on small problems) when division problems were solved via multiplication than when they were retrieved from memory, $F(1,57) = 16.69$, $MSe = 157848$. The strategy x load interaction showed larger effects of working-memory load (i.e., load RTs – no-load RTs) when division problems were solved via multiplication than when they were retrieved from memory, $F(1,57) = 5.05$, $MSe = 99248$.

There was also a significant interaction between working-memory component and load, $F(2,57) = 11.30$, $MSe = 102769$, which showed that load effects were significant for the executive component, $F(1,57) = 48.72$, $MSe = 102769$, but not for the active phonological component, $F(1,57) < 1$, or the passive phonological component, $F(1,57) = 2.19$ ($p = .14$). This interpretation was verified by separate ANOVAs that tested the effects of the different working-memory loads for each single strategy. Retrieval RTs were affected by executive loads, $F(1,57) = 75.27$, $MSe = 27985$ but not by active phonological or passive phonological loads (each $F < 1$). Via-multiplication RTs were affected by executive loads, $F(1,57) = 16.87$, $MSe = 174031$ but not by active phonological loads, $F(1,57) = 1.33$ ($p = .25$). However, via-multiplication RTs tended to be affected by passive phonological loads, $F(1,57) = 3.59$, $MSe = 174032$ ($p = .06$).

Table 6
Correlation table for Experiment 2 (division).

	Multiplication RT	Retrieval use[1]	Problem size	Arithmetic skill	Calculator use	Math experience	Math anxiety	Gender	Phon. passive	Phon. active	Exec.
Retrieval RT	.494*	-.149	.233*	-.264*	.019	-.047	.195*	-.130	-.020	-.014	.240*
Multiplication RT		-.206*	.210*	-.328*	.083	-.230*	.233*	-.105	.045	.027	.097
Retrieval use[1]			-.274*	.003	-.063	.150	-.006	.062	.041	.031	.000
Arithmetic skill					-.241*	.299*	-.321*	.002	--	--	--
Calculator use						.208*	.241*	-.030	--	--	--
Math experience							-.207*	.040	--	--	--
Math anxiety								-.128*	--	--	--
Gender									--	--	--

[1] Associative strength is operationalized by the participants' percentage retrieval use.
* $p < .0042$ (the Bonferroni-corrected *alpha* level of .05 when correlating 12 variables)
$df = 238$

To consolidate the results described above, and to investigate the influence of individual differences, correlations were calculated between strategy efficiency (i.e., retrieval RTs and via-multiplication RTs), strategy selection, working-memory load (i.e., executive, active phonological, and passive phonological), problem size, and individual-difference variables (i.e., math anxiety, arithmetic skill, calculator use, gender, and math experience). Correlation measures are presented in Table 6 (see also Footnote 6). Strategy efficiencies were smaller when problem size was higher and when the central executive was loaded, which confirms the ANOVA results. Strategy efficiencies correlated with several individual-difference variables as well. More specifically, retrieval and via-multiplication efficiencies were higher in high-skill participants than in low-skill participants, and higher in low-anxious participants than in high-anxious participants. Associative strength correlated significantly with the efficiency of the via-multiplication strategy but not with retrieval efficiency. Finally, the efficiency of the via-multiplication strategy was higher in more-experienced participants than in less-experienced participants.

Strategy selection. In order to investigate effects on strategy selection, a 3 x 2 x 2 ANOVA was conducted on percentages use of each single strategy (in the choice condition), with working-memory component (passive phonological, active phonological, executive) as between-subjects factor and load (no load vs. load) and size (small vs. large) as within-subjects factors (see Table 7).

For retrieval, the main effect of size was significant, $F(1,57) = 49.36$, $MSe = 10431$, indicating more frequent retrieval use on small problems (84%) than on large problems (71%). The main effects of load and working-memory component did not reach significance, and neither did any interaction (highest $F = 1.11$). The via-multiplciation strategy, in contrast, was used more frequently on large problems (29%) than on small problems (16%), $F(1,57) = 49.36$, $MSe = 10431$. None of the other effects reached significance (highest $F = 1.11$).

Table 7
Mean percentages strategy use as a function of load, working memory component[a], and size. Standard errors are shown between brackets.

		PHON passive		PHON active		Executive	
		No load	Load	No load	Load	No load	Load
Retrieval	Small	82 (5)	88 (5)	80 (5)	81 (5)	86 (5)	86 (5)
	Large	68 (6)	69 (5)	68 (6)	72 (5)	74 (6)	74 (5)
Via multiplic.	Small	18 (5)	12 (5)	20 (5)	19 (5)	14 (5)	14 (5)
	Large	32 (6)	31 (5)	32 (6)	28 (5)	26 (6)	26 (5)

[a] PHON = phonological.

In Table 6, the correlations between retrieval frequency, working-memory load, problem size, and individual differences are presented. Percentage retrieval use correlated with problem size but did not correlate with any of the working-memory loads, which confirms the ANOVA results. None of the individual-difference variables correlated significantly with strategy selection.

Secondary task performance. An analysis of variance was conducted on CRT accuracy, CRT speed, and letter-task accuracy (Table 4) with condition as within-subjects variable (single, choice, no-choice/retrieval, no-choice/via-multiplication). CRT accuracy differed across conditions, $F(3,17) = 11.80$, $MSe = 56$. More specifically, CRT accuracy was higher in the CRT-only condition (88%) than in the other conditions (75%), $F(1,19) = 33.86$, $MSe = 78$. CRT speed did not differ across conditions, $F(3,17) = 1.06$ ($p =$

.39). Performance on the active phonological task (i.e., the letter task) differed across conditions, $F(3,17) = 15.06$, $MSe = 180$. Accuracy was higher in the single-task condition (90%) than in dual-task conditions (72%), $F(1,19) = 13.26$, $MSe = 350$.

SUMMARY

Concerning *strategy efficiency*, it was shown that, as in Experiment 1, the role of the different working-memory resources differed across strategies. The retrieval strategy was affected by an executive load only, whereas the multiplication strategy was affected by an executive load and by a passive phonological load. Strategy efficiency further correlated significantly with several individual-difference variables; the interpretation of which is postponed to the general discussion. Also as in Experiment 1, *strategy selection* was not influenced by working-memory load.

GENERAL DISCUSSION

In the present study, the choice/no-choice method and the selective interference paradigm were combined in order to investigate the role of working memory in simple-arithmetic strategy selection and strategy efficiency. Results showed that the executive working-memory component was involved in all strategies (i.e., retrieval, transformation and counting in the multiplication experiment and retrieval and via-multiplication in the division experiment). Phonological working-memory components played a much smaller role, and tended to be needed in some nonretrieval strategies (i.e., counting in the multiplication experiment and via-multiplication in the division experiment).

THE ROLE OF EXECUTIVE WORKING-MEMORY RESOURCES

Executive working-memory resources were needed in direct retrieval of multiplication and division facts. Getting access to information stored in long-term memory is indeed one of the main executive (or attentional) functions (e.g., Baddeley, 1996; Baddeley & Logie, 1999; Cowan, 1995; Engle, Kane, & Tuholski, 1999a; Ericsson & Kintsch, 1995). Consequently, executive (or attentional) working-memory resources have for long been hypothesized to play a significant role in retrieving arithmetic facts from long-term memory (e.g., Ashcraft, 1992, 1995; Ashcraft, Donley, Halas, & Vakali, 1992; Barrouillet, Bernardin, & Camos, 2004; Geary & Widaman, 1992; Kaufmann, 2002; Kaufmann, Lochy, Drexler, & Semenza, 2003; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2000, 2002; Zbrodoff & Logan, 1986) and the present study succeeded to show this by using a rigorous method (i.e., solving simple-arithmetic problems in a no-choice/retrieval condition under an executive working-memory load).

We suppose that executive working-memory resources are needed to select the correct response. Indeed, the presentation of a simple multiplication or division problem does automatically activate several candidate answers in long-term memory (e.g., Campbell, 1997; De Brauwer & Fias, 2007; Galfano, Rusconi, & Umiltà, 2003; Rusconi, Galfano, Speriani, & Umiltà, 2004; Rusconi, Galfano, Rebonato, & Umiltà, 2006; Thibodeau, LeFevre, & Bisanz, 1996). After this automatic activation of several associated responses, a deliberate choice of the correct response has to be executed in order to complete the retrieval.

Executive working-memory resources did also play a role when nonretrieval strategies were used to solve multiplication or division problems. Of course, executing nonretrieval strategies does also require retrieval of known responses, which relies on executive resources. Moreover, executing nonretrieval strategies requires other demanding processes as well, such as performing calculations (e.g., Ashcraft, 1995;

Imbo, Vandierendonck, & De Rammelaere, in press e; Imbo, Vandierendonck, & Vergauwe, in press g; Logie, Gilhooly, & Wynn, 1994), manipulating interim results (Fürst & Hitch, 2000), and monitoring counting sequences (e.g., Ashcraft, 1995; Case, 1985; Hecht, 2002; Logie & Baddeley, 1987).

The central executive did not play a role in strategy selection: percentages of strategy use did not change under an executive working-memory load. This is in agreement with previous studies (e.g., Hecht, 2002; Imbo & Vandierendonck, in press b) and suggests that selecting simple-arithmetic strategies does not rely on executive working-memory resources. The absence of load effects on the strategy selection process is in agreement with the adaptive strategy choice model of Siegler and Shipley (1995). In this model, strategy selection is based solely on problem-answer association strengths (i.e., the answer that is most strongly associated with the presented problem is retrieved) and not on meta-cognitive processes such as executive (or attentional) processes.

THE ROLE OF PHONOLOGICAL WORKING-MEMORY RESOURCES

Phonological working-memory resources tended to be needed in nonretrieval strategies. More specifically, an active phonological load tended to affect the counting strategy in Experiment 1 (p = .10) and a passive phonological load tended to affect the via-multiplication strategy in Experiment 2 (p = .06). These results are in agreement with previous studies (Hecht, 2002; Imbo & Vandierendonck, in press b; Seyler et al., 2003) that also observed a significant role for the phonological loop in nonretrieval strategies.

The main function of the active phonological rehearsal process is storing intermediate and partial results (Ashcraft, 1995; Logie et al., 1994; Hitch, 1978), a function which is needed in nonretrieval strategies only. Without doubt, using the counting strategy to solve multiplication facts (e.g., 4 x 7 = 7… 14… 21… 28) requires storing intermediate results and thus relies on active phonological resources. The passive phonological store would come into play when more than one number needs to be maintained at any one time (Logie & Baddeley, 1987). This may explain why passive phonological resources were needed when the via-multiplication strategy was used to solve division problems. In order to transform a division problem into a multiplication problem (e.g., transforming 56 : 8 into 8 x ? = 56), participants have to maintain the dividend and the devisor while they are (sub-vocally) reciting their multiplication tables.

The present study also sheds further light on the equivocal results observed in previous studies investigating the role of the phonological loop in simple arithmetic. Whereas some studies did observe an effect of phonological load (e.g., Lee & Kang, 2002; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2002), others did not (e.g., De Rammelaere et al., 1999, 2001; Seitz & Schumann-Hengsteler, 2000). Present results suggest that strategy choices might have played a role. Studies in which participants relied more heavily on nonretrieval strategies might have observed larger effects of phonological working-memory loads than studies in which participants relied mainly on direct memory retrieval.

THE IMPACT OF INDIVIDUAL DIFFERENCES

Besides investigating the role of working memory in people's arithmetic strategy use, we also explored whether individual differences might influence strategy efficiency and/or strategy selection processes. In the following, the possible roles of these individual difference variables are discussed.

Arithmetic skill correlated significantly with all strategy efficiencies. More specifically, high-skill participants were more efficient (i.e., faster) in executing both retrieval and nonretrieval strategies to solve multiplication and division problems. This observation is not very surprising, since both the primary task (solving simple arithmetic problems) and the French Kit are speeded performance tests. Hence, correlations between arithmetic skill and strategy efficiency have been observed previously (e.g., Campbell & Xue, 2001; Imbo et al., in press f; Kirk & Ashcraft, 2001; LeFevre & Bisanz, 1986). Arithmetic skill correlated with strategy selection in the multiplication experiment only: high-skill participants used retrieval more frequently than did low-skill participants, an observation that is in agreement with previous studies as well (e.g., Imbo et al., in press f; LeFevre et al., 1996a, 1996b).

Associative strength (i.e., percentages retrieval use) correlated with retrieval efficiency in Experiment 1 but not in Experiment 2 (in which the correlation was quite high and in the correct direction, but not significant). Indeed, it has been asserted that problems with higher associative strengths are retrieved more efficiently from long-term memory (e.g., Ashcraft et al., 1992; Hecht, 2002). The correlation between associative strength and the via-multiplication strategy efficiency in Experiment 2 may be due to the fact that fast retrieval of multiplication facts is a critical component of this strategy.

Concerning *math anxiety*, the results of Experiment 1 indicated effects on strategy selection; retrieval use was significantly less frequent in high-anxious participants than in low-anxious participants. Anxious participants might set higher confidence criteria, which entails that they will only retrieve an answer when they are very sure about its correctness. No effects of math anxiety on strategy efficiency were found in Experiment 1, probably because solving simple multiplication problems is rather easy. Indeed, math anxiety would affect arithmetic performance only when the task is resource demanding (Ashcraft, 1995; Faust, Ashcraft, & Fleck, 1996). This reasoning also explains why math anxiety affected strategy efficiency in Experiment 2. In this experiment, in which division problems had to be solved, both retrieval and nonretrieval strategy use were less efficient in high-anxious participants than in low-anxious participants. Math-anxious participants are often occupied by worries and intrusive thoughts when performing arithmetic tasks. Because such thoughts load on working-memory resources, high-anxious participants have less working-memory resources left to solve the arithmetic task efficiently (Ashcraft & Kirk, 2001; Faust et al., 1996). It is reasonable that solving division problems is more resource-demanding than solving multiplication problems, which explains why math anxiety affected strategy efficiency in Experiment 2 but not in Experiment 1.

The frequency of *calculator use* correlated with strategy selection and strategy efficiency in Experiment 1 (multiplication) but not in Experiment 2 (division). More frequent calculator use was related to less efficient and less frequent retrieval use. Effects of calculator use on strategy efficiency have been observed earlier (Imbo et al., in press f), but no previous study observed a reliable effect of calculator use on simple-arithmetic strategy selection.

Math experience correlated with strategy selection and strategy efficiency. More-experienced participants used the retrieval strategy more frequently (Experiment 1) and were more efficient in the execution of the via-multiplication strategy (Experiment 2). Comparable effects have been observed previously (e.g., Imbo et al., in press f) and indicate that daily arithmetic practice has great effects on strategy selection and strategy efficiency.

Gender, finally, correlated with strategy selection and strategy efficiency in Experiment 1 but not in Experiment 2. When solving multiplication problems, men more frequently used retrieval than did women, an effect observed earlier (e.g., Carr & Jessup, 1997; Carr et al., 1999; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Geary et al., 2000b). We also observed more efficient retrieval use in men than in women, which confirms the hypothesis that gender differences in mental arithmetic are due to that fact that retrieval use is faster in men than in women (Royer et al., 1999a). However, gender might correlate with many other individual-difference variables such as calculator use, math experience, math anxiety and arithmetic skill. Hence, further research is needed to disentangle gender effects from other confounding variables.

Based on these exploratory correlations, it might be concluded that individual differences influence people's strategy efficiency and strategy selection processes. However, the effects were not always significant and differed across operations (multiplication vs. division) and across strategic performance measures (efficiency vs. selection). This was especially the case for the individual-difference variables which were based on one single question (e.g., calculator use, math anxiety). We acknowledge that the reliability of such measures can be questioned. Hence, future studies, in which individual differences are tested more thoroughly, are needed to confirm or disconfirm the exploratory results found here. For example, one might think to use the full Mathematics Anxiety Rating Scale (MARS, Richardson & Suinn, 1972) in order to test participants' math anxiety. Further research might also investigate the impact of individual differences in a more experimental way, e.g., by training participants, by manipulating their anxiety level, or by augmenting /reducing their calculator use.

CONCLUSION

The present study used a combination of two frequently used and approved methods, the selective interference paradigm and the choice/no-choice method. The selective interference paradigm enabled us to investigate the role of three different working-memory components; the choice/no-choice method enabled us to study strategy selection and strategy efficiency independently. Another novelty of the present study is that multiplication and division strategies were investigated. These operations differ greatly from addition and subtraction; already from childhood on up until adulthood. Moreover, the role of working memory in multiplication and division strategies has never been investigated before. A final novelty of the present study was that several individual-difference variables were included.

Concerning strategy efficiency, results showed that executive working-memory resources were involved in both retrieval and nonretrieval strategies. Active and passive phonological working-memory resources played a much smaller role and tended to be involved in nonretrieval strategies only. Strategy selection, in contrast, was not affected by executive or phonological working-memory loads. It was further shown that individual differences had a large impact as well. Arithmetic skill, calculator use, math experience, gender, and math anxiety influenced strategy efficiency and/or strategy selection. Individual differences should thus not be ignored when the cognitive systems underlying simple-arithmetic performance are investigated. Indeed, many effects caused by individual differences can be explained by cognitive variables. Effects of math anxiety for example, can be explained by working-memory limits (Ashcraft & Kirk, 2001; Faust et al., 1996) and effects of math experience can be explained by differential problem-answer strengths in long-term memory (Imbo et al., in press f). Arithmetic models and theories are challenged to incorporate these individual differences and their respective cognitive processes.

CHAPTER 6
EFFECTS OF PROBLEM SIZE, OPERATION, AND WORKING-MEMORY SPAN ON SIMPLE-ARITHMETIC STRATEGIES: DIFFERENCES BETWEEN CHILDREN AND ADULTS?

***Manuscript submitted for publication*[1,2]**

Adult's simple-arithmetic strategy use depends on problem-related characteristics, such as problem size and operation, and on individual-difference variables, such as working-memory span. The current study investigates (a) whether the effects of problem size, operation, and working-memory span on children's simple-arithmetic strategy use are equal to those observed in adults, and (b) how these effects emerge and change across age. To this end, simple-arithmetic performance measures and a working-memory span measure were obtained from 8-year-old, 10-year-old, and 12-year old children. Results showed that the problem-size effect in children results from the same strategic performance differences as in adults (i.e., size-related differences in strategy selection, retrieval efficiency, and procedural efficiency). Operand-related effects in children were equal to those observed in adults as well, with more frequent retrieval use on multiplication, more efficient strategy execution in addition, and more pronounced changes in multiplication. Finally, the advantage of having a large working-memory span was also present in children. The differences and similarities across children's and adult's strategic performance and the relevance of arithmetic models are discussed.

---

[1] This paper was co-authored by André Vandierendonck.

[2] Thanks are extended to the elementary school 'Gemeentelijke Basisschool' in Koksijde (Belgium), where all experiments were administered.

INTRODUCTION

The study of arithmetic performance is an important topic, since children spend a great deal of time mastering this skill and adults continue to use it in daily life. A well-documented observation is that a number of different strategies are used by adults as well as children to solve simple-arithmetic problems. Performance on a problem depends on both strategy selection and strategy efficiency. Strategy selection refers to the choice of a strategy among a set of alternatives available to solve the problem. In the domain of mental arithmetic, direct memory retrieval is distinguished from procedural strategies[3] such as counting (e.g., 7 + 4 = 7… 8… 9… 10… 11; 3 x 7 = 7… 14… 21) and transformation (e.g., 8 + 5 = 8 + 2 + 3; 9 x 6 = 10 x 6 – 6). Strategy efficiency refers to how fast and accurate strategies lead to the solution. Retrieval is generally more efficient than transformation, which is still more efficient than counting.

Accurate information about which strategies are applied (strategy selection) and how the strategies are applied (strategy efficiency) can be obtained by the combination of two approaches of data collection – self-reports and response latencies (Hopkins & Lawson, 2002). More precisely, in such a combined approach, trials are first separated by self-reports and then response latencies are analyzed. Using this combination of approaches, it has been shown that adult's strategic performance is influenced by both problem-related characteristics (such as operation and problem size) as well as by individual-difference variables (such as working-memory span). In the current study, we investigate whether or not children's strategic performance

[3] Many different labels have been used to denote what we call here 'procedural' strategies. Examples are "reconstructive strategies", "algorithmic strategies", "back-up strategies", "the usage of manipulatives", et cetera. In the current study, we consistently use the term 'procedural' strategies, to refer to (mostly time-consuming) strategies in which the solution is obtained in a sequence of operations.

is influenced by the same variables. Moreover, we also aimed at examining how the influence of these variables emerges and changes across primary school years, and whether these changes can be compared with practice or training effects in adults.

Effects of problem size on arithmetic strategy use. The problem-size effect, which refers to slower and more error-prone performance on large problems (e.g., 8 x 9) than on small problems (e.g., 2 x 3), is one of the most robust effects observed in mental-arithmetic research (Ashcraft, 1992; Zbrodoff, 1995). According to Campbell and Xue (2001), there are three strategy-related sources of the problem-size effect in adults: less frequent retrieval use for large than for small problems, lower retrieval efficiency for large than for small problems, and lower procedural efficiency for large than for small problems. In the current study, we investigated which of these sources determine the problem-size effect in children. More specifically, we checked for three different age groups (i.e., 8-, 10-, and 12-year olds) whether the problem-size effect was significant in terms of strategy selection (i.e., more frequent retrieval use on small than on large problems) and in terms of strategy efficiency (i.e., more efficient retrieval and procedural use on small than on large problems).

We also investigated whether the contribution of the different sources of the problem-size effect changes across the primary school years. Chronometric-only studies (i.e., without strategy reports) showed that the problem-size effect decreases gradually with age (e.g., Campbell & Graham, 1985; Cooney, Swanson, & Ladd, 1988; De Brauwer, Verguts, & Fias, 2006; Koshmider & Ashcraft, 1991). However, no study thus far used the combined approach (i.e., collecting self-reports and response latencies) to investigate which *strategic* sources contribute to the age-related decrease in the problem-size effect. We expected that age-related increases in retrieval use, retrieval efficiency, and procedural efficiency would be larger for large problems than for small problems (i.e., an age by size interaction).

Effects of operation on arithmetic strategy use. Adult studies consistently show operation-related differences in both strategy selection and strategy efficiency (e.g., Campbell, 1994; Campbell & Xue, 2001; Hecht, 1999; Imbo, Vandierendonck, & Rosseel, in press f). Generally, retrieval is used more frequently in multiplication than in addition, whereas both retrieval and procedural efficiencies are higher in addition than in multiplication. The joint investigation of children's performance in addition and multiplication is rather scarce, however. Lépine, Roussel, and Fayol (2003) investigated 5th graders' addition and multiplication verification performance (e.g., 2 + 3 = 7, true/false?). Although children did not have to report which strategies they used, Lépine et al. (2003) used priming techniques to infer which strategies the children used. Based on the observation that priming the operation sign (+ or x) reduced addition response times but not multiplication response times, they inferred that addition problems were generally solved by means of procedures whereas multiplication problems were rather solved by direct fact retrieval. To test whether this operation-dependent effect on strategy selection changes across age, Lépine et al. (2003) compared their results with those obtained by Roussel, Fayol, and Barrouillet (2002), who tested the same verification problems in adults. Apart from faster response times in adults than in 5th graders, similar effects occurred in both age groups. Consequently, Lépine et al. (2003) conclude that addition and multiplication problems are solved similarly by 5th graders and adults, i.e., by means of procedural and retrieval strategies, respectively.

The current study aimed to test (a) at what age these operation-related differences originate, and (b) whether or not these operation-related differences change across age. In contrast to Lépine et al. (2003), who used a verification task without strategy reports, we used a production task with trial-by-trial strategy reports. Moreover, we tested three different age groups (2nd, 4th, and 6th graders) whereas Lépine et al. tested 5th graders only. The retrieval bias for multiplication over addition was expected to originate from 2nd grade on. Indeed, as addition is taught already in 1st grade, children in the current study (2nd, 4th, and 6th graders) should master this operation reasonably well. As multiplication is taught in 2nd grade, the youngest children in the current study were only starting to master this operation. Because multiplication performance is strongly based on direct fact retrieval, we expected larger increases in retrieval use for multiplication than for addition, and especially between 2nd and 4th grade. From 4th grade on, we expected to observe the same operation-related differences in children as in adults; i.e., more frequent retrieval use in multiplication than in addition and more efficient strategy execution in addition than in multiplication.

Effects of working-memory span on arithmetic strategy use. It has been shown that working memory, a memory system involved in concurrent maintenance and processing of information (Baddeley, 1996; Baddeley & Logie, 1999), plays a significant role in adults' arithmetic performance (see DeStefano & LeFevre, 2004, for a review on dual-task studies). Low-span adults have been shown to perform worse on arithmetic tasks than high-span adults (e.g., Jurden, 1995; Seyler, Kirk, & Ashcraft, 2003), but it is not known whether this effect is due to individual differences in strategy selection, strategy efficiency, or both. Working memory in children has been studied in relation to mathematical disabilities (see Geary, 2004, for a review) rather than in relation to its role in normally developing children. As respects strategy selection, higher working-memory spans have been linked with less frequent use of procedural strategies and more frequent use of retrieval strategies (e.g., Barrouillet & Lépine, 2005; Geary, Bow-Thomas, Liu, & Siegler, 1996a; Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Noël, Seron, & Trovarelli, 2004; Steel & Funnel, 2001). Working memory has also been related to strategy efficiency. Adams and Hitch (1997) observed faster arithmetic performance in children with higher working-memory spans. Since no strategy reports were obtained, it is not clear whether working-memory span was correlated with both retrieval and procedural efficiency. Rasmussen and Bisanz (2005) observed that several working-memory measures together accounted for a substantial proportion of the variability in arithmetic performance ($R^2 \geq .40$). Comparable results were obtained by

Swanson (2004; Swanson & Beebe-Frankenberger, 2004), who showed that working memory predicted solution accuracy of arithmetic word problems independent of other skill measures such as fluid intelligence, reading skill, math skill, and short-term memory. Comparably, Noël et al. (2004) observed that children's addition accuracy was not predicted by processing speed but that it was predicted by several measures of working-memory capacity. Finally, Barrouillet and Lépine (2005) observed that direct memory retrieval was faster in high-span children than in low-span children.

The current study investigated (a) whether working-memory capacity is differently correlated with children's strategy selection and strategy efficiency, and (b) whether the importance of having a large working-memory capacity stays equally important throughout the primary school years. Several predictions were made. First, since poor working-memory resources may result into weaker and less accessible associations in long-term memory, we predict more frequent retrieval use in high-span children than in low-span children. Second, since poor working-memory resources may lead to smaller amounts of available attentional resources, we predict more efficient retrieval use and more efficient procedural use in high-span children than in low-span children. Indeed, attentional resources are needed to activate items in long-term memory and to maintain this activation (Barrouillet, Bernardin, & Camos, 2004; Lovett, Reder, & Lebiere, 1999) and to execute several subprocesses needed in procedural strategies (DeStefano & LeFevre, 2004). Finally, we hypothesized that the advantage of having a large working-memory capacity would decrease with age. The latter prediction was based on previous findings arguing that cognitive resources are most important during the initial phase of skill acquisition whereas their role declines as facts become represented in long-term memory (e.g., Ackerman, 1988).

Simple-arithmetic models. Finally, we also wanted to test whether arithmetic models are able to explain our results. One model that is especially relevant in the present context is the adaptive strategy choice model (ASCM) of Siegler and Shipley (1995). In this model, people have several strategies available. When encountered with a simple-arithmetic problem, they will try to choose the fastest and most accurate strategy among all available strategies. However, people also set a confidence criterion, which determines how sure they must be to state a retrieved answer, and a search length, which determines how many attempts they will make to retrieve an answer before trying a procedural strategy to solve the problem. One will thus retrieve the answer from long-term memory only if the problem can be solved fast and accurately with the retrieval strategy. Otherwise stated, the retrieval time may not exceed the search length criterion whereas the confidence criterion should be exceeded. If the retrieval strategy would provide a slow and/or incorrect answer (e.g., when a problem is associated with several possible answers in long-term memory) and thus exceeds the search length or does not exceed the confidence criterion, one will rather use a procedural strategy to solve the problem.

The ASCM also predicts the efficiency with which retrieval and procedural strategies will be executed. The efficiency of retrieval strategies depends on the number of searches in long-term memory. If the distribution of problem-answer associations is peaked (i.e., only one answer receives high activation), the correct answer will be retrieved very fast. If the distribution of problem-answer associations is flat (i.e., many answers receive activation), more time is needed to search the correct answer among several incorrect (but highly related) answers. The efficiency of procedural strategies, in contrast, does not depend on the peakedness of problem-answer associations but on the difficulty of executing the particular procedural strategy. For example, the number of counts determines the efficiency of the counting strategy. In the current study, we explicitly tested this prediction of Siegler's model. More specifically, because both retrieval frequency and retrieval efficiency rely on the peakedness of problem-answer associations, it was hypothesized that retrieval frequency would be highly correlated with retrieval efficiency. However, because procedural efficiency does not rely on the distribution of problem-answer associations but rather on the number of steps to be executed, procedural frequency, or its component, retrieval frequency, should not be correlated with procedural efficiency.

The present study. To summarize, the purpose of the present study was to test whether the effects of problem size, operation, and working-memory span observed in children are similar to those observed in adults. We also wanted to test whether age-related effects in children can be compared with practice effects in adults. Several predictions were made. Since the magnitude of the problem-size effect decreases with age, we expected that the age-related increase in retrieval use and strategy efficiency would be larger for large problems than for small problems (i.e., an age by size interaction). We further expected larger increases in retrieval use and strategy efficiency for multiplication than for addition, and especially between 2nd and 4th grade (i.e., an age by operation interaction). Finally, we hypothesized that having a large working-memory capacity would correlate with more frequent retrieval use and more efficient strategy execution. However, the advantages of having a large working-memory capacity were expected to decrease with age.

METHOD

PARTICIPANTS

Sixty children participated. They all attended the same elementary school in the Flemish part of Belgium. Twenty of them were in the 2nd grade of elementary school (mean age: 8 years 0 months; 9 girls and 11 boys), twenty other children were in the 4th grade of elementary school (mean age: 10 years 0 months; 10 girls and 10 boys), and the last twenty children were in the 6th grade of elementary school (mean age: 12 years 0 months; 7 girls and 13 boys). The children were selected from the whole ability range, although those who were considered by their teachers to have specific learning or behavioral difficulties were excluded. The children had no documented brain injury, socio-cultural disadvantage, or behavioral problems. The children only participated when they, as well as their teachers and their parents consented.

To verify whether the three age groups were representative samples of the population, a standardized skill test [Arithmetic Tempo Test (ATT), De Vos, 1992] was administered. This pen-and-paper test consists of several subtests that require very elementary computations (e.g., 2 + 3 = ?). Each subtest concerns only one arithmetic operation (addition, subtraction, multiplication, or division). In the present experiment, the first two subtests were administered, i.e., the addition and the subtraction subtest, each consisting of 40 items of increasing difficulty. We opted for these operations since the problems of the multiplication and division subtest were beyond the 2nd graders' skill (e.g., 12 x 4, 75 : 25). The children were given 1 minute for each subtest and had to solve as many problems as possible within that minute. Performance on the test was the sum of both subtests. An ANOVA on these performance data with grade (2, 4, 6) as between-subjects factor showed a main effect of grade, $F(2,57) = 61.6$, with increasing performance across the 2nd, 4th, and 6th grade (scores of 29.4, 48.3, and 55.9, respectively). We further tested whether the children differed from the expected ATT performance, a normal development considered. Therefore, the score expected at the moment of testing (i.e., 19 educational months for 2nd graders, 39 educational months for 4th graders, and 59 educational months for 6th graders) was compared with each child's individual score. Paired-samples *t*-tests (two-tailed) showed no significant differences between observed and expected ATT performance, with *t* values of 0.4, 1.9, and 1.7 for 2nd, 4th, and 6th graders respectively (all *p*s > .05). Clearly, the three age groups were representative subgroups of the population.

MATERIALS AND PROCEDURE

The children were individually tested in the month of May. At that moment, even the 2nd graders had learned how to solve simple addition problems (up to 20) and simple multiplication problems (up to 100). All children were administered a simple-arithmetic task in which they had to solve addition and multiplication problems, and a reading-span task to test their complex working-memory span.

Simple-arithmetic task. All children solved 56 simple addition problems and 56 simple multiplication problems. The problems were constructed from all the possible pair-wise combinations of the integers 2 to 9 with tie problems (e.g., 2 + 2, 2 x 2) excluded. For both addition and multiplication problems, small problems were defined as problems with a product smaller than 25 and large problems as problems with a product larger than 25. The order of operation was counterbalanced for all grades. For 2nd graders only, the addition and multiplication test were administered on two consecutive days, so as to keep the total session load manageable. For 4th and 6th graders both operations succeeded each other immediately. For each operation, five practice trials were presented to let the children get used to the task and the material.

The problems were presented one at a time in the centre of a computer screen. A trial started with a fixation point for 500 msec. Then the problem was presented horizontally in Arabic format as dark-blue characters on a light-grey background, with the operation sign (+ or x) at the fixation point. Children were asked to verbally state their answer as soon as they knew it. The problem remained on screen until the child responded. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, children wore a microphone which was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 msec. The experimenter entered the answer by means of the numerical path of the keyboard. On each trial, visual feedback was presented to the children, a happy face (smiley) when their answer was correct, and a sad face when their answer was incorrect.

Immediately after having solved the problem, children were asked to report verbally which strategy they had used to solve the problem. Taking into account the literature on strategy use in simple arithmetic, a distinction was made between three levels of strategies, namely Retrieval, Transformation, and Counting. Retrieval was explained as "remembering or knowing the answer directly from memory". If children said that the answer "just popped into their head", their strategy was coded as retrieval. Transformation was explained as "deriving the answer from some known facts" Examples were given, such as making an intermediate step to 10 (e.g., 8 + 5 = 8 + 2 + 3; 9 x 4 = 10 x 4 - 4) and using a tie in order to solve a non-tie problem (e.g., 6 + 7 = 6 + 6 + 1; 5 x 6 = 5 x 5 + 5). Counting was explained as "step-by-step counting to get the answer. For addition, this meant counting one-by-one, e.g., 4 + 3 = 4… 5… 6… 7. No distinction was made between finger counting, counting all, or counting from the larger addend. For multiplication, counting meant (subvocally) reciting the multiplication tables, e.g., 8 x 3 = 8… 16… 24. A fourth category 'Other' was added to cover the case when the children used another strategy or did not know which strategy they had used. All incorrect trials and all trials that were corrupted due to failure of the voice activated relay were repeated at the end of the block to decrease the amount of data loss.

Reading-span task. This is a *working-memory* span task, in which materials (i.e. words) have to be maintained in memory while other information (i.e., sentences for comprehension) has to be processed. This task differs from *short-term memory* span tasks in which small amounts of materials have to be maintained and recalled without any processing load being imposed. The reading-span task (Daneman & Carpenter, 1980) is a classical example of a working-memory span task. In this task, participants have to read sets of increasing numbers of sentences aloud while retaining words in memory. The Dutch reading-span task used in the current study (see also De Jonge & De Jong, 1996) included two practice trials (consisting of two sentences) and two trials for each consecutive number of sentences (range: 2 – 7 sentences). The sentences were presented on a sheet of paper one-by-one and the child had to read them aloud. At the end of each sentence, a word was given by the experimenter, which had to be stored in memory. At the end of each sentence set, the child had to reproduce all words in the order in which they had been presented by the experimenter. As the number of sentences increased, the number of words-to-retain increased as well. For example, if the child had read 4 sentences after which each time a word was provided, the correct response after having read all the sentences consisted of 4 words. If the child failed at remembering the words in two sets with the same number of sentences/words, the reading-span task was stopped. The score on this task was the number of correctly remembered words (range: 4 – 54 words).

RESULTS

Of all trials, 13% was spoiled due to failure of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failure of the sound-activated relay to 3%. Further, all trials on which children had used a strategy of the 'Other' category (0.3%) were deleted. Finally, all the response times (RTs) more than 3 standard deviations from each participant's mean (per operation) were discarded as outliers (2%).

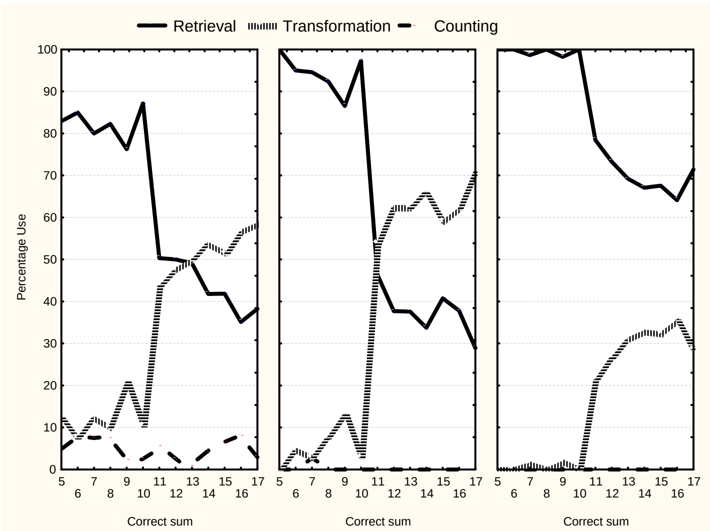
The results section is divided into two parts. First, analyses of variance (ANOVAs) were carried out to investigate age-related effects on strategy selection and strategy efficiency; problem size and operation taken into account. Second, regression analyses were performed to test whether working-memory span plays a role in children's simple-arithmetic strategy use. All reported results are considered to be significant if $p < .05$, unless mentioned otherwise.

ANALYSES OF VARIANCE

Strategy selection. The children used several strategies to solve the simple-arithmetic problems. Not all strategies were chosen equally frequently, however. As can be seen in the upper panel of Figure 1, children of all grades most often chose retrieval to solve addition problems with sums smaller than 10. As soon as the sum was larger than 10, a tremendous decrease in retrieval use was observed. The 2nd and 4th graders chose the transformation strategy more often than direct memory retrieval on problems with a solution above 10. For the 6th graders, in contrast, retrieval was the most frequently used strategy to solve addition problems with sums both smaller and larger than 10. Analyses on the subject level showed that only eleven children used retrieval on all addition problems, two 2nd graders, one 4th grader and eight 6th graders.

The lower panel of Figure 1 shows the strategy choice pattern for multiplication problems. Here, retrieval use decreased linearly with increasing problem size, whereas the frequency of transformation increased as problem size became larger. Another striking difference with the strategy choices for addition problems is that direct memory retrieval was the most popular strategy already from 2nd grade on and on all problem sizes. Analyses on the subject level showed that fourteen children used retrieval on all multiplication problems, one 2nd grader, two 4th graders and eleven 6th graders. The very infrequent use of counting (in both addition and multiplication) was probably due to the curriculum in Belgium, which strongly advices children against using this strategy.

Addition



Multiplication

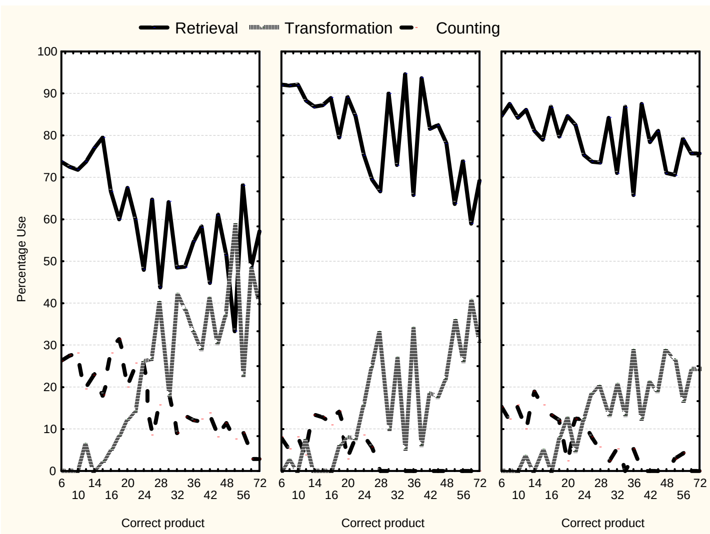


Figure 1
Percentages of the used strategies for each grade (left: 2nd grade, middle: 4th grade, right: 6th grade), as a function of operation and problem size.

A 3 (Grade: 2, 4, 6) x 2 (Problem size: small vs. large) x 2 (Operation: addition vs. multiplication) ANOVA with repeated measures on problem size and operation was conducted on percentages of retrieval use (see Table 1). The main effect of grade was significant, $F(2,57) = 4.6$, $MS_e = 8625$. The percentage of retrieval use increased linearly across 2^{nd}, 4^{th} and 6^{th} grade, $F(1,57) = 9.2$. We first tested whether this effect differed across small and large problems. Obviously, retrieval was more frequently used on small than on large problems, $F(1,57) = 94.6$, $MS_e = 299$. Although this effect of problem size was true in all grades [$F(1,57) = 28.7$, 58.9, and 14.5, for 2^{nd}, 4^{th}, and 6^{th} grade, respectively], problem size interacted with grade, $F(2,57) = 3.8$. The age-related trends in retrieval use thus differed as a function of problem size. More specifically, the increase in retrieval use between 2^{nd} and 4^{th} grade was only significant for small problems, $F(1,57) = 7.9$, and not for large problems, $F(1,57) < 1$, whereas the increase in retrieval use between 4^{th} and 6^{th} grade was only significant for large problems, $F(1,57) = 3.9$, and not for small problems, $F(1,57) < 1$. Consequently, the difference in retrieval use between small and large problems (i.e., % retrieval use on large problems – % retrieval use on small problems) did not decrease between 2^{nd} and 4^{th} grade ($p > .10$), but did decrease significantly between 4^{th} and 6^{th} grade, $F(1,57) = 7.5$.

We then tested whether the increase in retrieval use differed across operations. Although the main effect of Operation did not reach significance, $F(1,57) = 1.6$, the operation x grade interaction did, $F(2,57) = 5.1$. Between 2^{nd} and 4^{th} grade, there was a significant increase in retrieval use for multiplication problems, $F(1,57) = 6.1$, and not for addition problems, $F(1,57) < 1$, whereas the reverse was true between 4^{th} and 6^{th} grade, with a significant increase in retrieval use for addition problems, $F(1,57) = 8.6$, and not for multiplication problems, $F(1,57) < 1$. This age-related pattern also explains why percentages of retrieval use did not differ across addition and multiplication problems in 2^{nd} graders, $F(1,57) < 1$: Retrieval was used as frequently for both operations. Percentages retrieval use did not differ across operations in 6th graders either, with very frequent retrieval use on both addition problems and multiplication problems, $F(1,57) < 1$. In the 4th grade, however, percentages retrieval use were larger on multiplication problems than on addition problems, $F(1,57) = 11.1$.

Table 1
Mean percentages retrieval use for 2nd, 4th, and 6th graders as a function of operation and problem size. Standard errors are shown between brackets.

		2nd Grade	4th Grade	6th Grade
Addition	Small	79 (3)	88 (3)	97 (3)
	Large	45 (8)	37 (8)	71 (8)
Multiplication	Small	64 (6)	85 (6)	81 (6)
	Large	56 (6)	76 (6)	78 (6)

Retrieval efficiency: latency. A 3 (Grade: 2, 4, 6) x 2 (Problem size: small vs. large) x 2 (Operation: addition vs. multiplication) ANOVA with repeated measures on problem size and operation was conducted on correct retrieval RTs[4] (see Table 2). The main effect of grade was significant, $F(2,57) = 65.2$, $MS_e = 1380585$. Retrieval RTs linearly decreased as children were older, as confirmed by a planned comparison with linear contrast, $F(2,57) = 113.3$. We first tested whether this increase in retrieval efficiency differed between small and large problems. The main effect of problem size

[4] Since (a) not all children used both retrieval and procedural strategies, and (b) only RTs of the correctly solved problems were analyzed, for some children empty cells occurred in the ANOVAs on latencies. We replaced these empty cells for each child with the correct RT of the corresponding cell [i.e., the mean RT (over participants) of the grade x problem size x operation cell].

was significant, $F(1,57) = 116.6$, $MS_e = 225284$, with faster retrieval on small than on large problems. Moreover, the retrieval problem-size effect (retrieval RTs on large problems – retrieval RTs on small problems) was significant in all grades, $F(1,57) = 146.2$, 15.7, and 7.0 for 2nd, 4th and 6th grade, respectively. Grade interacted with problem size, however, $F(2,57) = 26.1$. This interaction showed that the retrieval problem-size effect decreased significantly between 2nd and 4th grade $F(1,57) = 33.0$, but did not differ between 4th and 6th grade, $F(1,57) < 1$. More precisely, between 2nd and 4th grade, children became faster in retrieving large problems, whereas between 4th and 6th grade, they became slightly faster in retrieving both small and large problems.

We then tested whether the increase in retrieval efficiency differed across operations. Retrieval was faster on addition problems than on multiplication problems, $F(1,57) = 6.2$, $MS_e = 448058$, but operation did not interact with grade. The age-related increase in retrieval efficiency thus runs parallel for addition and multiplication.

Table 2
Mean retrieval latencies (msec) for 2nd, 4th, and 6th graders as a function of operation and problem size. Standard errors are shown between brackets.

		2nd Grade	4th Grade	6th Grade
Addition	Small	2252 (113)	1195 (113)	966 (113)
	Large	3806 (261)	1682 (261)	1328 (261)
Multiplication	Small	2844 (130)	1454 (130)	1177 (130)
	Large	3856 (142)	1809 (142)	1378 (142)

Retrieval efficiency: accuracy. An Arcsin transformation was applied to the proportions of correct solutions. The same 3 x 2 x 2 ANOVA was conducted on these Arcsin transformed values. To enhance comprehension, however, Table 3 depicts percentages of correct answers. The main effect of grade did not reach significance, $F(2,57) = 1.9$, but planned comparisons showed a significant increase in retrieval accuracy between 2nd and 4th grade, $F(1,57) = 3.6$, and no difference between 4th and 6th grade, $F(1,57) < 1$. Small problems were retrieved significantly more accurately than large problems, $F(1,57) = 9.8$, and this effect did not change across age [i.e., no grade x problem size interaction, $F(2,57) < 1$]. There was also a trend towards higher accuracies on addition problems than on multiplication problems, $F(1,57) = 3.3$ with $p = .07$, but this effect did not change across age either [i.e., no grade x operation interaction, $F(2,57) < 1$].

Table 3
Mean retrieval accuracies (%) for 2nd, 4th, and 6th graders as a function of operation and problem size. Standard errors are shown between brackets.

		2nd Grade	4th Grade	6th Grade
Addition	Small	98 (1)	99 (1)	100 (1)
	Large	95 (1)	98 (1)	98 (1)
Multiplication	Small	95 (1)	99 (1)	98 (1)
	Large	96 (1)	97 (1)	95 (1)

Procedural efficiency: latency. The same 3 x 2 x 2 ANOVA was conducted on correct procedural RTs (see Table 4 and footnote 4). The main effect of grade was significant, $F(2,57) = 29.8$, $MS_e = 6988988$. The 2nd graders were significantly slower than 4th and 6th graders, $F(1,57) = 53.0$ and $F(1,57) = 34.2$, respectively, whereas there was no difference between 4th and 6th graders, $F(1,57) = 2.0$. We first tested whether this age-related effect differed across small and large problems. Obviously, procedures were executed faster on small than on large problems, $F(1,57) = 38.9$, MS_e = 1093316. Problem size interacted with grade, though, $F(2,57) = 18.3$: the procedural problem-size effect (procedural RTs on large problems – procedural RTs on small problems) decreased significantly between 2nd and 4th grade children, $F(1,57) = 31.9$, but did not differ between 4th and 6th grade children, $F(1,57) < 1$. Between 2nd and 4th grade, children became faster in executing procedures on large problems, which reduced the problem-size effect. Consequently, the procedural problem-size effect was significant in 2nd grade, $F(1,57) = 71.9$, but not in 4th and 6th grade.

We also tested whether the increase in procedural efficiency differed across operations. The main effect of operation was significant, with higher procedural efficiencies on addition problems than on multiplication problems, $F(1,57) = 52.5$, MS_e = 3820986. Operation also interacted with grade, $F(2,57) = 3.0$. Planned comparisons showed that the increase in procedural efficiency was larger for multiplication problems than for addition problems between 2nd and 4th grade, $F(1,57) = 5.3$, but did not differ across operations between 4th and 6th grade, $F(1,57) < 1$.

Table 4
Mean procedural latencies (msec) for 2nd, 4th, and 6th graders as a function of operation and problem size. Standard errors are shown between brackets.

		2nd Grade	4th Grade	6th Grade
Addition	Small	3172 (248)	1547 (248)	2096 (248)
	Large	5089 (265)	2027 (265)	2446 (265)
Multiplication	Small	5786 (548)	3139 (548)	3525 (548)
	Large	7854 (484)	2939 (484)	4033 (484)

Procedural efficiency: accuracy. The same 3 x 2 x 2 ANOVA was conducted on the Arcsin transformations of proportions of correct answers. Percentages of correct answers are shown in Table 5. The main effect of grade was significant, $F(2,57) = 13.5$. A planned comparison confirmed that procedural accuracies increased linearly with grade, $F(1,57) = 24.7$. We first tested whether this age-related effect differed across small and large problems. Accuracies were higher on small than on large problems, $F(1,57) = 9.8$. Furthermore, problem size interacted with grade, $F(2,57) = 4.5$. Whereas the procedural problem-size effect (accuracy on large problems – accuracy on small problems) decreased significantly between 2nd and 4th grade, $F(1,57) = 8.4$, it did not change anymore between 4th and 6th grade, $F(1,57) < 1$. Consequently, the procedural problem-size effect was significant in 2nd grade, $F(1,57) = 17.2$ but not in 4th and 6th grade (both Fs < 1.5).

Table 5
Mean procedural accuracies (%) for 2nd, 4th, and 6th graders as a function of operation and problem size. Standard errors are shown between brackets.

		2nd Grade	4th Grade	6th Grade
Addition	Small	99 (1)	95 (1)	100 (1)
	Large	87 (3)	98 (3)	100 (3)
Multiplication	Small	86 (2)	98 (2)	99 (2)
	Large	78 (4)	95 (4)	97 (4)

We then tested whether the increase in procedural efficiency differed across operations. Accuracies were higher on addition problems than on multiplication problems, $F(1,57) = 14.6$. Operation also interacted with grade, $F(2,57) = 3.1$. Addition accuracies did not increase between 2nd and

4^{th} grade, but did increase between 4^{th} and 6^{th} grade, $F(1,57) = 4.8$. Multiplication accuracies, in contrast, increased between 2^{nd} and 4^{th} grade, $F(1,57) = 15.7$, but not between 4^{th} and 6^{th} grade, $F(1,57) < 1$.

Summary. Analyses of variance were run to answer two questions formulated in the introduction. Concerning the first question (Are all three sources of the problem-size effect present in children and do they change across age?), results showed that two sources of the problem-size effect were present in all grades, namely more frequent retrieval use on small than on large problems, and more efficient retrieval use on small than on large problems. The third source of the problem-size effect, more efficient procedural use on small than on large problems, was only present in the 2^{nd} grade. Moreover, the different sources of the problem-size effect changed across age: The decrease in the size of the problem-size effect was first (i.e., between 2^{nd} and 4^{th} grade) due to more efficient retrieval use and more efficient procedural use, after which (i.e., between 4^{th} and 6^{th} grade) it was due to more frequent retrieval use.

Concerning the second question (Does children's simple-arithmetic strategy use differ between addition and multiplication, and does this difference change across age?), results showed that the improvement in strategic performance (more frequent retrieval use and more efficient procedural use) on multiplication was especially apparent between 2^{nd} and 4^{th} grade. The improvement in strategic performance on addition, in contrast, was especially apparent between 4^{th} and 6^{th} grade. Finally, the age-related improvement in retrieval efficiency was equally large for addition and multiplication.

In the next section, regression analyses are run to test the role of working-memory span in children's simple-arithmetic strategy use across the primary school years (cf. our third research question).

REGRESSION ANALYSES

Before presenting the results of the regression analyses, we first report children's performance on the working-memory span task, which is one of the predictors used in the regression analyses. Children's performance on the working-memory span task was tested with an ANOVA with grade (2, 4, 6) as between-subjects factor. As expected, the main effect of grade was significant, $F(2,57) = 37.8$, which indicates an increasing working-memory span across the 2nd, 4th, and 6th grade (scores of 12.7, 20.2, and 28.0, corresponding to working-memory spans of 2.4, 3.3, and 4.2, respectively). The question now is whether working-memory span plays a role in strategy selection or strategy efficiency. More precisely, we will test whether working-memory span predicts variance in percentages of retrieval use, retrieval latencies, and procedural latencies. To this end, correlation and regression analyses were conducted for each dependent variable (i.e., retrieval use, retrieval latency, and procedural latency[5]) separately.

Retrieval use was regressed on working-memory span, problem size[6], and operation. Retrieval latency and procedural latency were regressed on the same three variables and on percentage retrieval use as well. Doing so, we wanted to test Siegler's (1988b) prediction that retrieval frequency should correlate with retrieval efficiency but not with procedural efficiency. Indeed, both retrieval frequency and retrieval efficiency are in his model of strategy choice dependent on the peakedness of problem-answer

[5] As age-related differences were substantially smaller in accuracy data than in latency data (cf. ANOVA results), regression analyses were performed on latency data only.

[6] In the regression analyses, problem size was determined by the correct answer of the problem (i.e., sizes from 5 to 17 for addition problems and sizes from 6 to 72 for multiplication problems). Thus, whereas a dichotomous measure of problem size was used in the analyses of variance, a continuous measure of problem size was used in the regression analyses. Operation was in the regression analyses coded by a dummy variable with value 1 for addition problems and value -1 for multiplication problems.

associations, whereas procedural efficiency depends on the difficulty of executing the particular procedure. Correlation results can be found in Table 6 and regression results can be found in Table 7.

Retrieval use. The percentage of retrieval use was regressed on problem size, operation, and working-memory span, for each grade separately (see Table 7). For 2nd graders, the total amount of variance explained (R^2) was .073, $F(3,778) = 20.36$. Retrieval use was significantly more frequent on small than on large problems, and for high-span children than for low-span children. Operation was not significantly predictive. For the 4th grade, $R^2 = .088$ and $F(3,1056) = 34.12$, smaller problem sizes predicted more frequent retrieval use, whereas working-memory span did not play a significant role. Direct memory retrieval was also more frequently used on multiplication problems than on addition problems. For the 6th grade finally, $R^2 = .013$ and $F(3,975) = 4.45$, retrieval use was more frequent on small than on large problems, whereas no effects of operation or working-memory span were observed. Note that operation was only significantly predictive of percentage retrieval use in 4th graders, which fits well with the ANOVA on retrieval use.

Retrieval efficiency. Retrieval latencies were regressed on problem size, operation, working-memory span, and percentage retrieval use, for each grade separately (see Table 7). The R^2 was .192 for the 2nd grade, $F(4,585) = 34.73$, $R^2 = .226$ for the 4th grade, $F(4,819) = 59.77$ and $R^2 = .162$ for the 6th grade, $F(4,818) = 39.66$. Problem size was significantly predictive in all grades, with faster retrieval use on small than on large problems. The percentage retrieval use was significantly predictive in all grades as well, with faster retrieval use when retrieval was more frequently used. Working-memory span was only predictive in 2nd and 4th grade, with faster retrieval use for high-span than for low-span children. Operation was not predictive in any grade.

Table 6
Correlations between retrieval use, retrieval RTs, procedural RTs, problem size, operation, and working-memory span.

2nd Grade	Retrieval RT	Procedural RT	Problem size	Operation	WM span
Retrieval use	-.239*	-.164*	-.138*	.017	.218*
Retrieval RT		.445*	.344*	-.159*	-.174*
Procedural RT			.293*	-.163*	-.284*
4th Grade	**Retrieval RT**	**Procedural RT**	**Problem size**	**Operation**	**WM span**
Retrieval use	-.164*	.057	-.018	-.224*	-.045
Retrieval RT		.547*	.402*	-.212*	-.224*
Procedural RT			.342*	-.315*	-.285*
6th Grade	**Retrieval RT**	**Procedural RT**	**Problem size**	**Operation**	**WM span**
Retrieval use	-.320*	-.132	-.102*	.021	.022
Retrieval RT		.303*	.251*	-.116*	-.035
Procedural RT			.412*	-.253*	.228*

* *p* is significant at the 0.01 level (2-tailed)

Table 7
Summary of the regression analyses for variables predicting percentage retrieval use, retrieval RTs, and procedural RTs.

Retrieval use	2nd Grade			4th Grade			6th Grade		
	B	*SE*	*ß*	*B*	*SE*	*ß*	*B*	*SE*	*ß*
Problem size	-0.50	0.11	-.194*	-0.64	0.10	-.237*	-0.33	0.09	-.140*
Operation	-3.48	1.70	-.087	-15.15	1.52	-.366*	-2.33	1.47	-.063
WM span	2.03	0.32	.218*	-0.31	0.20	-.044	0.12	0.18	.022
Retrieval RTs	2nd Grade			4th Grade			6th Grade		
	B	*SE*	*ß*	*B*	*SE*	*ß*	*B*	*SE*	*ß*
Problem size	47.74	6.19	.354*	19.03	1.88	.388*	9.64	1.42	.270*
Operation	42.00	93.21	.021	-2.36	29.70	-.003	24.87	21.79	.045
WM span	-75.13	18.55	-.153*	-27.68	3.91	-.218*	-0.54	2.61	-.007
Retrieval use	-15.82	3.07	-.196*	-4.95	1.24	-.125*	-13.22	1.36	-.312*
Procedural RTs	2nd Grade			4th Grade			6th Grade		
	B	*SE*	*ß*	*B*	*SE*	*ß*	*B*	*SE*	*ß*
Problem size	86.27	15.38	.302*	22.65	5.66	.252*	39.43	7.32	.387*
Operation	198.2	237.4	-.045	-195.8	92.04	-.134	-94.96	116.1	-.059
WM span	-302.3	46.47	-.272*	-69.36	11.01	-.284*	63.38	15.73	.241*
Retrieval use	-14.52	6.57	-.094	-6.86	2.47	-.126	-3.51	3.41	-.062

* *p* is significant at the 0.01 level

Procedural efficiency. Procedural latencies were regressed on problem size, operation, working-memory span, and percentage retrieval use, for each grade separately (see Table 7). The R^2 was .174 for the 2nd grade, $F(4,475) = 25.03$, $R^2 = .216$ for the 4th grade, $F(4,392) = 27.04$ and $R^2 = .238$ for the 6th grade, $F(4,227) = 17.71$. Problem size was significantly predictive in all grades, with more efficient procedural strategy execution on small than on large problems. Working-memory span was predictive in all grades as well, but the relation between span and procedural efficiency changed across grades. High-span 2nd and 4th graders were more efficient procedural strategy users than were low-span 2nd and 4th graders, but high-span 6th graders were *less* efficient procedural strategy users than were low-span 6th graders. Span-related differences in strategy selection can explain this unexpected result. Indeed, retrieval was used more frequently in high-span 6th graders (85%) than in low-span 6th graders (78%), and this difference was larger for large problems (79% vs. 69%) than for small problems (91% vs. 87%). Consequently, high-span 6th graders used procedural strategies to solve the largest problems only, which results in large procedural RTs in high-span 6th graders. Indeed, procedural RTs for large problems were larger for high-span than for low-span 6th graders (3707 msec vs. 3089 msec, respectively), whereas procedural RTs for small problems did not differ between high-span and low-span 6th graders (2772 msec vs. 2533 msec, respectively). Operation and percentage retrieval did not predict procedural efficiency in any grade.

Summary. The advantage of having a large working-memory span decreased across grades, especially regarding retrieval frequency and retrieval efficiency. More specifically: (1) working-memory span significantly predicted retrieval frequency for 2nd graders but not for 4th and 6th graders, and (2) working-memory span predicted retrieval efficiency for 2nd and 4th graders but not for 6th graders. Comparably, the execution of procedural strategies benefited from a high working-memory span in 2nd and 4th grade only. Because high-span 6th graders used procedural strategies almost exclusively on large problems, procedural efficiency *de*creased for these children. Strategy efficiency data were thus influenced by the children's strategy choices. This bias can be avoided by using the choice/no-choice method (Siegler & Lemaire, 1997), as discussed further in this paper. A final interesting observation was that, as predicted by Siegler (1988b), percentage retrieval use did predict retrieval efficiency but not procedural efficiency.

GENERAL DISCUSSION

Children's arithmetic strategic performance increased with age: older children used memory retrieval more often, were faster and more accurate in retrieving arithmetic facts, and were faster and more accurate in executing procedural strategies. In the remaining of this chapter, we discuss whether or not (and from which moment on) children's arithmetic strategy use resembles adults' arithmetic strategy use. We successively discuss the problem-size effect, operation-related effects, and the role of working-memory span. The discussion section ends with an evaluation of the present results within a model of arithmetic strategic performance.

THE PROBLEM-SIZE EFFECT

From 4^{th} to 6^{th} grade, the frequency of retrieval use increased for large problems but not for small problems. As such, the problem-size effect caused by *strategy selection* processes became smaller from 4^{th} grade on. However, the retrieval strategy was more frequently used on small than on large problems in all age groups. Strategy selection processes were thus a significant source of the problem-size effect in 2^{nd}, 4^{th}, and 6^{th} graders. From 2^{nd} to 4^{th} grade, retrieval and procedural efficiencies increased for large problems but not for small problems. This way, the problem-size effect caused by *strategy efficiency* processes became smaller from 2^{nd} grade on. However, the retrieval problem-size effect (i.e., the difference in retrieval RTs between large and small problems) stayed significant in all grades. The procedural problem-size effect (i.e., the difference in procedural RTs between large and small problems), in contrast, was significant in 2nd grade only. Thus, whereas problem size affected retrieval frequency and retrieval efficiency in all age groups, problem size affected procedural efficiency in 2nd grade only.

Importantly, previous studies showed that all three size-related effects on strategy use (i.e., less frequent retrieval use for large than for small problems, lower retrieval efficiency for large than for small problems, and lower procedural efficiency for large than for small problems) are significant sources of the problem size effect in adults (e.g., Campbell & Xue, 2001). The first two factors were significant in the current child study as well. Concerning the third factor, something strange occurred: the procedural problem-size effect was present in 2nd grade but disappeared in 4th and 6th grade. However, it reappeared in secondary-school children (Imbo et al., in press f). The proficiency for solving small and large problems equally efficiently by means of procedures is probably caused by practice and schooling effects. As soon as children finish elementary school, such effects disappear, resulting in less efficient procedure execution, especially for large problems. Comparable effects have been reported by Geary (1996), who observed that the problem-size effect disappeared and reversed between 1st and 3rd grade Chinese children, but re-appeared in Chinese adults. The investigation of the appearance, disappearance, and re-appearance of the problem-size effect across lifetime provides interesting ideas for future research.

To conclude, the decreasing problem-size effect was associated with an increase in strategy efficiency for younger children and with an increase in retrieval frequency for older children. Moreover, the size-related effect on strategy efficiency did not change anymore from 4th grade on. Since De Brauwer et al. (2006) observed that the problem-size effect remains equally large from 6th grade on till adulthood; we might conclude that children from 4th grade on have developed a memory network that strongly resembles the adult memory network. This conclusion is in agreement with previous studies which maintain that mental-arithmetic networks might be completely operational from 3rd grade on (e.g., Ashcraft & Fierman, 1982; Koshmider & Ashcraft, 1991; Lemaire, Barrett, Fayol, & Abdi, 1994).

DIFFERENCES BETWEEN ADDITION AND MULTIPLICATION

Children of all age groups retrieved addition facts more efficiently (i.e., faster and more accurately) than they retrieved multiplication facts. Comparable effects have been observed in adults (e.g. Campbell & Xue, 2001; Hecht, 1999; Imbo et al., in press f). Importantly, the main effect of operation in *retrieval efficiency* (i.e., more efficient retrieval use for addition than for multiplication) did not change across the primary school years. This indicates consistent changes in the speed with which addition and multiplication facts are retrieved from long-term memory. Thus, although the addition and multiplication network may differ across age groups in general (i.e., main differences in retrieval speed), their development seems to run fairly parallel (i.e., no interaction between age and operation). Otherwise stated, addition and multiplication involve similar retrieval processes across childhood (this study) and in adulthood (e.g., Campbell & Oliphant, 1992; Geary, Widaman, & Little, 1986; Miller, Perlmutter, & Keating, 1984).

Another persistent effect in adults is that retrieval is used more frequently in multiplication than in addition (e.g. Campbell & Xue, 2001; Hecht, 1999; Imbo et al., in press f). Surprisingly, in the current study, this operation-related effect on *strategy selection* was significant in 4th grade only. In 2nd grade, where children only start to learn the multiplication tables, retrieval was used as frequently in addition as in multiplication. However, multiplication fact retrieval significantly increased from 2nd to 4th grade, an effect that was probably due to the great emphasis of the Belgian school system on the memorization of multiplication tables. Because the amount of retrieval use in multiplication reached in 4th grade (81%) was comparable to that observed in Belgian adults (73% - 88%; Imbo et al., in press f; Imbo &

Vandierendonck, 2007a, in press c) and in North-American adults (64% - 88%, Campbell & Xue, 2001), there was no space left for another increase in retrieval use. Thus, for multiplication, children from 4^{th} grade on may already have developed a complete memory network that strongly resembles an adult network (see also De Brauwer et al., 2006). For addition, retrieval use still increased from 4^{th} to 6^{th} grade. This effect should be attributed to general practice effects rather than to specific training effects.

Finally, *procedural efficiency* was higher in addition than in multiplication, an effect observed in adults as well (e.g. Campbell & Xue, 2001; Hecht, 1999; Imbo et al., in press f). All age groups showed higher efficiencies in addition than in multiplication and this effect boosted from 4^{th} grade on. Indeed, between 2^{nd} and 4^{th} grade, procedural efficiency increased for multiplication but not for addition. This early increase in multiplication efficiency might be related to two facts. First, multiplication is heavily trained from 2^{nd} grade on. As noted above, children are taught to solve multiplication problems very fast and accurately. Obviously, increases in procedural efficiency are the precursors of increases in retrieval use. Second, multiplication strategy efficiency is more susceptible to change than addition strategy efficiency. In a previous study, we observed that adults' daily arithmetic experience (e.g., the amount of arithmetic lessons in secondary school) influenced multiplication efficiency but not addition efficiency (Imbo et al., in press f). Moreover, when explicitly practicing arithmetic problems, effects were larger in multiplication efficiency than in addition efficiency (Imbo & Vandierendonck, 2007a). These effects can be accounted for by the relative efficiency of procedural strategies for addition and multiplication. Indeed, in both children and adults, addition strategies are far more efficient (i.e., faster) than multiplication strategies. Consequently, it is less demanding to increase multiplication efficiencies than to increase addition efficiencies. Strategy selection and strategy efficiency processes in other arithmetic operations (such as subtraction and division) have been investigated less frequently (but see Campbell & Xue, 2001; Imbo & Vandierendonck, in press b,c; Robinson et al., 2006; Seyler et al., 2003) and are an issue for further research.

THE ROLE OF WORKING MEMORY

Working-memory span was related to *strategy selection* in 2nd grade only. High-span 2nd graders used retrieval more frequently than did low-span 2nd graders, but this effect disappeared in 4th and 6th graders. Up until now, the relevance of working memory in children's strategy selection process was equivocal, since some studies observed a correlation between working-memory span and retrieval use (e.g., Barrouillet & Lépine, 2005), whereas others did not (e.g., Geary et al., 2004). The current study suggests that strategy selection processes tend to rely on working-memory resources in young children only. This runs parallel to adult studies which observed that working memory is not needed in strategy selection (e.g., Hecht, 2002; Imbo & Vandierendonck, in press b,c).

Working-memory span was related to *retrieval efficiency* in 2nd and 4th grade, with less efficient retrieval use in low-span children than in high-span children. Less efficient retrieval use in low-span children than in high-span children has been observed earlier (e.g., Barrouillet & Lépine, 2005). Apparently, low-span children develop flatter distributions of problem-answer associations, resulting in less frequent and less efficient retrieval use. However, with growing age and increasing practice and schooling, even low-span children can develop peaked distributions of problem-answer associations. Consequently, the differences between low-span children and high-span children decrease across primary school years. In adults, the role of working-memory in strategy efficiency has been investigated by means of dual-task studies rather than correlational studies. The evidence is equivocal: Hecht (2002) observed no effects of working-memory load on retrieval efficiency whereas Imbo and Vandierendonck (in press b,c) did observe less efficient fact retrieval under working-memory load. Future research is needed to specify the relation between working memory and direct fact retrieval in both children and adults.

Finally, *procedural efficiency* was also related to working-memory span. In 2nd and 4th grade, low-span children executed procedural strategies less efficiently than did high-span children. A significant role of working memory in procedural strategy execution has been observed in adults as well (e.g., Hecht, 2002; Imbo et al., in press b,c). The role of working memory in procedural strategies is quite obvious: Each procedure requires several subprocesses that require working-memory resources, such as storing intermediate results, keeping track of several steps, integrating information, et cetera (see DeStefano & LeFevre, 2004, for a review). High-span children and adults can carry out these various subprocesses with fewer demands on a limited resource pool than low-span children and adults. Consequently, high-span children have more working-memory resources left for storage while processing the problem, resulting in higher procedural efficiency scores.

Surprisingly, we observed higher procedural efficiencies in low-span 6th graders than in high-span 6th graders. We suppose that this effect was due to an artifact. Indeed, high-span 6th graders used procedural strategies to solve the largest problems only, which might have increased their procedural RTs relative to low-span 6th graders, who used procedural strategies on smaller problems as well. This artifact occurred because we only used a choice condition, in which strategy efficiency data are biased by strategy selection effects (Siegler & Lemaire, 1997). Such a bias would have been avoided by using the choice/no-choice method (devised by Siegler & Lemaire, 1997), which not only entails a choice condition, but also no-choice conditions. In no-choice conditions, participants are asked to use one single strategy to solve all problems. In a recent study using the choice/no-choice method, we indeed showed that loading 6th graders' working-memory resources resulted in less efficient procedural strategy use (Imbo & Vandierendonck, in press d).

Finally, it should be noted that the advantage of having a large working-memory capacity decreased across age. The relation between working-memory span and retrieval frequency was not significant anymore from 4th grade on, and the relation between working-memory span and retrieval efficiency was not significant anymore from 6th grade on. These results are consistent with Ackerman's (1988) findings. Specifically, working memory is most important during the initial phase of arithmetic-skill acquisition and its role declines as procedures are used less frequently and facts become represented in long-term memory. Working-memory resources might thus be needed to achieve a complete representation of number facts in long-term memory (e.g., Geary, 1990; Geary & Brown, 1991; Hitch & McAuley, 1991; Siegler & Shrager, 1984), which explains the correlation between working-memory span and retrieval use in the younger children. However, once the number facts are completely represented in long-term memory, fact retrieval becomes more automatic and less effortful, resulting in smaller arithmetic-performance differences between high-span children and low-span children.

A MODEL OF STRATEGIC CHANGE

In the ASCM (Siegler & Shipley, 1995), outlined in the introduction, people have several strategies available and try to choose the best one. Strategy selection occurs on the basis of knowledge on each strategy's efficiency. Each time a simple-arithmetic problem is solved correctly, the problem-answer association increases, resulting in a more peaked distribution of problem-answer associations. The more peaked the distribution of problem-answer associations, the more frequently retrieval is used; while the use of procedural strategies vanishes. This reasoning fits with our data, since the frequency of retrieval use increased across age.

Across age, the efficiency of both retrieval and procedural strategies increased as well. This observation can also be accounted for by the ASCM.

Indeed, each time an answer is retrieved from long-term memory, the problem-answer association is strengthened. As outlined above, this results in more peaked distributions of problem-answer associations and thus in more efficient retrieval use. Each execution of a procedural strategy brings an increase in the strategy's speed and a decrease in its probability of generating an error. The ASCM thus accounts for age-related increases in both retrieval efficiency and procedural efficiency.

To summarize, in the ASCM, both retrieval frequency and retrieval efficiency depend on the peakedness of the problem-answer association, whereas procedural efficiency does not. We tested the hypothesis that the frequency of retrieval use would be highly correlated with retrieval efficiency but not with procedural efficiency. Regression results confirmed this prediction, since the frequency of retrieval use was highly predictive of retrieval efficiency in all grades, whereas the frequency of retrieval use was not predictive of procedural efficiency.

In the following, we verify whether the ASCM is able to account for the size-related, operation-related, and resource-related results observed in children and adults. First, what does the ASCM tell about the problem-size effect? Because small problems are more frequently encountered, young children develop peaked problem-answer associations for small problems and relatively flat problem-answer associations for large problems. They might also set larger search lengths for small problems, because they are taught that small problems should be retrieved in any case. More peaked problem-answer associations and larger search lengths for small problems than for large problems results in increases in retrieval frequency for small problems but not in increases in retrieval efficiency for small problems, which is exactly what we observed between 2nd and 4th grade. However, as children grow older, the emphasis shifts towards large problems. This results in more peaked problem-answer associations for large problems and thus in more frequent retrieval use for large problems, as observed between 4th and 6th grade.

However, some results are more difficult to explain by the ASCM. Theoretically, the ASCM predicts that extensive practice should create equally peaked problem-answer associations for small and large problems. Accordingly, retrieval frequency and retrieval efficiency should be equal in small as in large problems. Such effects, however, have not yet been observed; the retrieval problem-size effect is still present in adults (e.g., Campbell & Xue, 2001; Imbo et al., in press f; LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b), even after explicit practice (e.g., Imbo & Vandierendonck, 2007a). One explanation for the persisting problem-size effect is that adult's problem-answer associations are still stronger for small problems than for large problems. Another explanation is based on interference effects (Campbell, 1987a, 1995): answers to large problems undergo stronger interference effects than do answers to small problems, resulting in less efficient retrieval for large than for small problems. Future research is needed to clear out the persistent nature of the problem-size effect across lifetime.

Second, what can the ASCM tell about the differences across addition and multiplication? More frequent retrieval use for multiplication than for addition suggests more peaked distributions of associations for multiplication than for addition. However, multiplication facts are retrieved less efficiently than addition facts, which suggests rather the opposite (i.e., more peaked distributions for addition than for multiplication). Thus, although the notion of peaked and flat distributions is relevant *within* each operation (i.e., peaked distributions predict frequent and fast retrieval use), it is not clear how this notion can account for differences in retrieval frequency and retrieval speed *across* operations. One possibility is that the search length is larger for multiplication than for addition. Indeed, this would result in more frequent retrieval use for multiplication than for addition and in less efficient (i.e., slower) retrieval use for multiplication than for addition. The more efficient procedural use for addition than for multiplication can easily be accounted for by the ASCM, in which procedural efficiencies are based on the amount and the difficulty of steps. Because fewer increments of counting steps are needed in addition procedures than in multiplication procedures, procedural strategies are much easier to implement in addition than in multiplication.

Finally, we consider whether the ASCM may account for the role of working memory in strategic performance. Although working memory is not explicitly included in the ASCM, it has been predicted (e.g., Geary, 1993, 1994) that low-span children use slow counting procedures, which lead to delays in problem encoding and consequently to weak problem-answer associations and flat distributions of associations. High-span children, in contrast, develop strong problem-answer associations and more peaked distributions of associations, resulting in more frequent and more efficient retrieval use. More frequent and more efficient retrieval use in high-span children than in low-span children is exactly what we observed. However, with growing age and experience, even low-span children may develop strong problem-answer associations and more peaked distributions of associations. Hence, the differences in retrieval performance between low-span and high-span children should decrease across the primary school years, as was observed in the current study.

It should be noted that other models may also account for the relation between working-memory span and retrieval performance. The time-based resource-sharing model of working memory (Barrouillet et al., 2004), for example, predicts that lower working-memory resources reduce the amount of attentional resources available to activate knowledge from long-term memory (see also Cowan, 1999; Lovett et al., 1999). Consequently, poor working-memory resources not only impair the formation of associations in long-term memory but also the retrieval of existing associations. In other models (e.g., Engle, 2001; Engle, Kane, & Tuholski, 1999a), lower working-memory resources reduce the ability to resist interference, which might also result in less efficient retrieval performance. Thus far, our and other's results (e.g., Barrouillet & Lépine, 2005) do not contradict the theories mentioned above. Future research is thus needed to investigate the predictive value of several models in the domain of mental arithmetic and cognitive strategy use, and more specifically to investigate the specific role of working memory in retrieval frequency and retrieval efficiency.

ADDENDUM

As noted in the introduction (Chapter 1), the original results section of the study reported in Chapter 6 included state trace analyses as well. Although these analyses did not reach publication, we do report them here. The advantage of state trace analyses is that they control for general speeding effects. Such control is especially needed in developmental research, since the faster performance observed in older children may lead in itself to smaller problem-size effects.

Hence, in this additional results section, we further tested the problem-size effect observed in retrieval and procedural latencies. Note that the standard analyses of variance reported in Chapter 6 do not differentiate between general differences in response speed and specific age-related changes. More specifically, as younger children are generally slower than older children, such speeding differences may have accounted for the larger problem-size effect in younger children than in older children. Indeed, as a matter of fact, slower latencies implicate larger effect sizes (e.g., Verhaeghen & Cerella, 2002). One method that is able to correct for age-group differences in general processing speed is the state trace method. An advantage of this method is that it does not assume additive or multiplicative effects (see also De Brauwer et al., 2006). Indeed, as one does not know *a priori* whether age and experimental effects are additively or multiplicatively related, results obtained by methods relying on additive and/or multiplicative assumptions (e.g., analyses of variance, *z*-scores, logarithmic and ratio transformations, analyses of proportional slowing) cannot always be justified. Note that the state trace analyses relied on exactly the same data sets as used for the analyses reported previously in Chapter 6.

STATE TRACE ANALYSES

For the state trace analyses, mean response time (RT) data on large problems were plotted against mean RT data on small problems. This was done for retrieval trials on additions, for retrieval trials on multiplications, for procedural trials on additions, and for procedural trials on multiplications separately. We thus obtained four scatter plots (or 'state spaces') in which each point corresponded to one subject of whom the mean RT on small problems was plotted against the mean RT on large problems.

Regression analyses were then performed to quantitatively describe the problem-size effect. If a different problem-size effect between two age groups is only influenced by a general scaling effect of increased processing speed, one single regression line suffices to explain the relation between small and large problems in these two age groups. Different regression lines, in contrast, indicate that there is a true difference between two age groups concerning the problem-size effect. To test whether or not one single line suffices to explain the problem-size effect in different age groups, RTs on large problems were regressed on three predictor variables: RTs on small problems, a dummy variable that codes for the contrast between two age groups (e.g., -1 for 2nd graders, 1 for 4th graders, and 0 for 6th graders to test 2nd versus 4th graders), and a variable that is the product of the RTs on small problems and the corresponding contrast dummy variable[7]. We then tested whether the restricted regression model (with RTs on small problems as the only predictor) was as good as the full regression model (with all three predictors). If both models do not significantly differ from each other, general speeding effects suffice to explain the larger problem-size effects observed in younger children. If the full model is significantly better than the restricted model, there are specific age-related changes in the magnitude of

---

[7] For a more statistical elaboration of this method, we refer to De Brauwer et al. (2006).

the problem-size effect. In the following, we only report the *F*-change values (i.e., full model – restricted model).

Retrieval problem-size effect. We first tested whether the problem-size effect differed between 2nd and 4th grade. The *F*-change model was not significant for addition, $F < 1$, but it was for multiplication, $F(2,56) = 4.6$, indicating different problem-size effects in 2nd and 4th grade. Between 4th and 6th grade, the *F*-change model was significant neither for addition nor for multiplication (both *F*s < 1), indicating that the retrieval problem-size effect did not change anymore after 4th grade. Obviously, between 2nd and 6th graders, the *F*-change model was significant for multiplication, $F(2,56) = 10.0$, but not for addition, $F < 1$. Hence, one single line fitted the retrieval problem-size effect for 2nd, 4th, and 6th graders in addition problems (see Figure 1a). For the retrieval problem-size effect in multiplication problems, two lines were needed: one for the 2nd graders and one for the 4th and 6th graders (see Figure 1b). Note that each line displays the mean RTs on large problems as a function of the mean RTs on small problems for a number of age groups that do not differ from each other in their problem-size effect.

Procedural problem-size effect. When comparing 2nd and 4th graders' performance on additions solved with procedures, the *F*-change model reached significance, $F(2,56) = 8.2$, whereas it did not between 4th and 6th grade, $F < 1$. For multiplication a comparable pattern was observed, with a significant *F*-change model between 2nd and 4th grade, $F(2,56) = 21.3$ but not between 4th and 6th grade, $F < 1$. Obviously, between 2nd and 6th graders, the *F*-change model was significant for addition, $F(2,56) = 11.1$ and for multiplication, $F(2,56) = 11.1$. Hence two lines were needed to fit the procedural problem-size effect in addition and multiplication: one for the 2nd graders and one for the 4th and 6th graders (see Figures 2a and 2b).

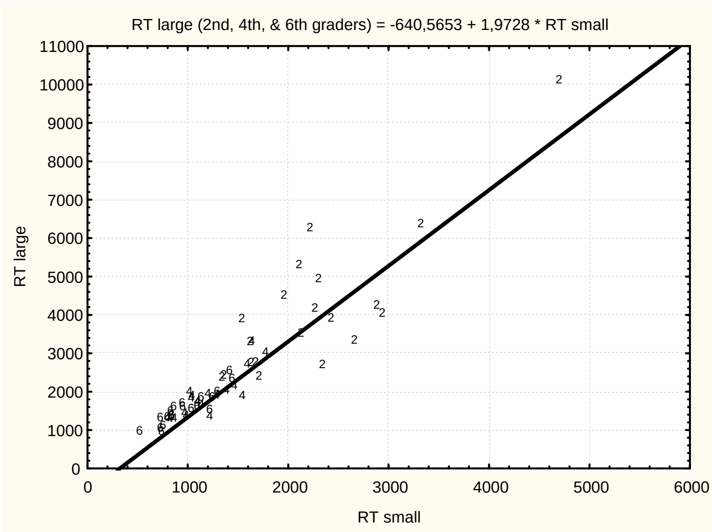


Figure 1a
State trace of the retrieval problem-size effect in addition. The text labels (2, 4, or 6) point to the grade to which the corresponding observation belongs. As there was no difference between 2nd, 4th, and 6th graders, there is only one line.

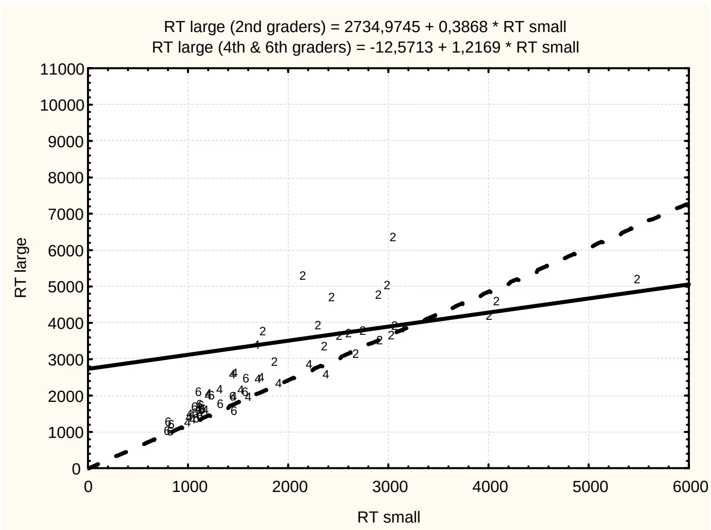


Figure 1b
State trace of the retrieval problem-size effect in multiplication. The text labels (2, 4, or 6) point to the grade to which the corresponding observation belongs. The black line is the fit for the 2nd graders, whereas the dotted line is the fit for the 4th and 6th graders.

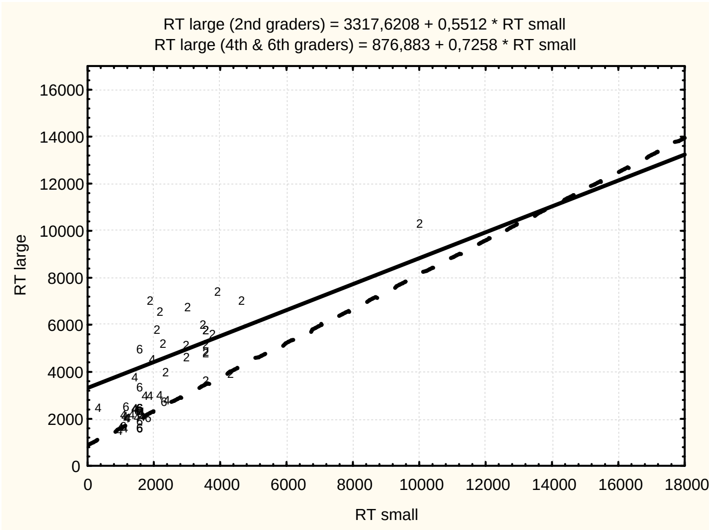


Figure 2a
State trace of the procedural problem-size effect in addition. The text labels (2, 4, or 6) point to the grade to which the corresponding observation belongs. The black line is the fit for the 2nd graders, whereas the dotted line is the fit for the 4th and 6th graders.

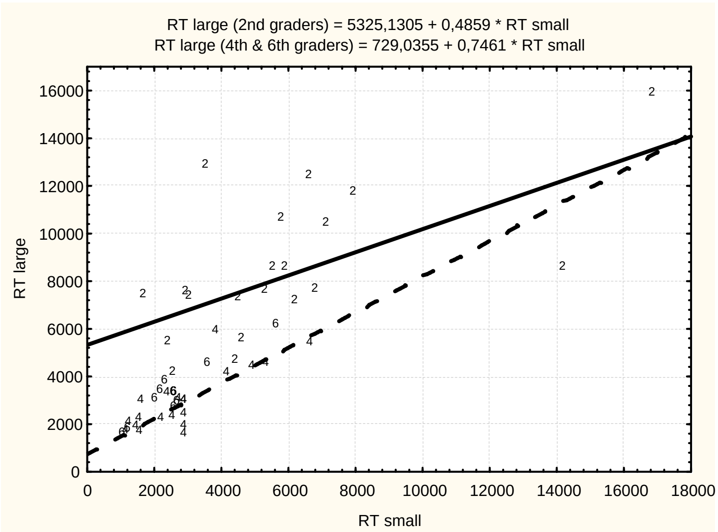


Figure 2b
State trace of the procedural problem-size effect in multiplication. The text labels (2, 4, or 6) point to the grade to which the corresponding observation belongs. The black line is the fit for the 2nd graders, whereas the dotted line is the fit for the 4th and 6th graders.

DISCUSSION

The results of the state trace analyses confirmed what was stated in the analyses of variance reported in Chapter 6. More specifically, the decreasing problem-size effect between 2nd graders and 4th graders was specifically age-related and was not due to general speeding differences across age groups. The state trace analyses further showed that neither the retrieval problem-size effect nor the procedural problem-size effect decreased anymore after 4th grade, although there might still be general increases in processing speed. Otherwise stated, when general speeding is taken into account, children from 4th grade on show a latency advantage for small over large problems that stays equal across development. Since De Brauwer et al. (2006) observed that the problem-size effect remains equally large from 6th grade on till adulthood, we might conclude that children from 4th grade on have developed a memory network that strongly resembles an adult memory network. This conclusion is in agreement with previous studies which maintain that addition and multiplication networks might be completely operational from 3rd grade on (e.g., Ashcraft & Fierman, 1982; Koshmider & Ashcraft, 1991; Lemaire et al., 1994).

It should be noted though, that De Brauwer et al. (2006) did observe a decrease in the problem-size effect between 4th and 6th grade whereas we did not. This inconsistency between both studies may result from a different operationalization of problem size. Whereas we defined small problems as problems with a product smaller than 25 and large problems as problems with a product larger than 25, De Brauwer et al. (2006) defined small problems as problems with both operands smaller than 5 and large problems as problems with both operands larger than 5. As ties were excluded, this indicates that the small stimuli of De Brauwer et al. (2006) consisted of only three problems (i.e., 2 x 3, 2 x 4, and 3 x 4). Future studies are needed to investigate whether the range of small and large problems influences the magnitude of the problem-size effect *in se* and across development.

CHAPTER 7
THE DEVELOPMENT OF STRATEGY USE IN ELEMENTARY-SCHOOL CHILDREN: WORKING MEMORY AND INDIVIDUAL DIFFERENCES

Journal of Experimental Child Psychology (in press)[1,2]

The current study tested the development of working-memory involvement in children's arithmetic strategy selection and strategy efficiency. To this end, an experiment – in which the dual-task method and the choice/no-choice method were combined – was administered to 10-, 11-, and 12-year-old children. Working memory was needed in retrieval, transformation, and counting strategies, but the ratio between available working-memory resources and arithmetic task demands changed across age. More frequent retrieval use, more efficient memory retrieval, and more efficient counting processes reduced the working-memory requirements. Strategy efficiency and strategy selection were also modified by individual differences such as processing speed, arithmetic skill, gender, and math anxiety. Short-term memory capacity, in contrast, was not related to children's strategy selection or strategy efficiency.

---

[1] This paper was co-authored with André Vandierendonck

[2] Thanks are extended to the elementary school 'St. Lievens – Kolegem' in Mariakerke (Belgium), where all experiments were administered.

INTRODUCTION

Learning to perform simple-arithmetic tasks efficiently and with little effort is one of the most fundamental skills taught during the elementary-school years. Several cognitive mechanisms may underpin the development of arithmetic skill in children. The current study was designed to investigate the role of one such cognitive mechanism, namely the executive component of working memory. Besides an online study of the role of working memory in the development of children's arithmetic strategy use, we tested the influence of individual-difference variables such as processing speed, short-term memory, arithmetic skill, math anxiety, and gender.

THE ROLE OF WORKING MEMORY IN CHILDREN'S ARITHMETIC STRATEGY USE

Working memory can be defined as a set of processing resources of limited capacity involved in information maintenance and processing (e.g., Baddeley & Logie, 1999; Engle, Tuholski, Lauglin, & Conway, 1999; Miyake, 2001). Most researchers agree that working-memory resources play a role in children's simple-arithmetic performance. This assertion is mainly based on studies showing a working-memory deficit in mathematically disabled children (e.g., Bull, Johnston, & Roy, 1999; Geary, Hoard, & Hamson, 1999; McLean & Hitch, 1999; Passolungi, Cornoldi, & De Liberto, 1999; Passolunghi & Siegel, 2004; Swanson & Beebe-Frankenberger, 2004; Swanson & Sachse-Lee, 2001; van der Sluis, de Jong, & van der Leij, 2004). The goal of our study, however, was to investigate the role of working memory in arithmetic strategy use by normally developing children. To this end, we needed to overcome several shortcomings of the studies just mentioned.

First, the role of working memory has been studied predominantly by means of *correlations* between working-memory measures (e.g., counting span, Trails task, Stroop task) and simple-arithmetic performance (e.g., Bull & Scerif, 2001; Bull & Johnston, 1997; Bull et al., 1999; McLean & Hitch, 1999; Passolunghi & Siegel, 2001). Because correlation is not causation, it is still possible that working-memory measures and mathematical ability rely on a common factor such as general intelligence or processing speed.

In the current study, we aimed at investigating the role of working memory in children's arithmetic performance *online*. To this end, we used the dual-task method, in which children needed to solve simple-arithmetic problems (i.e., the primary task) while their working memories were loaded by means of the secondary task. The dual-task method has been used frequently in adult studies (for a review, see DeStefano & LeFevre, 2004), which clearly show that working memory is needed in adults' simple-arithmetic performance. More specifically, adults' simple-arithmetic performance always relies on *executive* working-memory resources, as opposed to verbal and visuo-spatial working-memory resources, of which the role in simple arithmetic is less clear.

Although the dual-task method has been used only rarely in child studies, Hitch, Cundick, Haughey, Pugh, and Wright (1987) conducted a dual-task study in which children had to verify simple addition problems (e.g., 3 + 5 = 7, true/false?) while their memories were phonologically loaded. Because errors and latencies rose under such a load, Hitch and colleagues concluded that children's counting processes involve inner speech. The dual-task method was further used by Kaye, deWinstanley, Chen, and Bonnefil (1989). In their study, 2^{nd}, 4^{th}, and 6^{th} graders verified simple addition problems while their working memories were loaded by means of a probe detection task. This secondary task affected addition speed most profoundly in 2^{nd} graders, and much less so in 4^{th} and 6^{th} graders, indicating that computational efficiency increases with increasing grade level. Finally, Adams and Hitch (1997) did not use the dual-task method but rather manipulated the presentation format of addition problems (i.e., oral vs. visual presentation). The visual presentation provided an external record of the addends that reduced working-memory load. Because children's performance was better in the visual condition than in the oral condition, Adams and Hitch concluded that children's mental-arithmetic performance is mediated by working-memory resources. Unfortunately, none of these studies investigated the impact of an executive working-memory load on children's arithmetic performance.

A second shortcoming in previous studies is the ignorance of the locus of effect of working-memory support. Although it has been shown that working-memory resources correlate with arithmetic performance, it is not clear whether working memory is needed in strategy selection processes (i.e., which strategies are chosen to solve the problem?) and/or strategy efficiency processes (i.e., is the problem solved fast and accurately by means of the chosen strategy?). This is a relevant question, however, because children do use several strategies to solve simple-arithmetic problems (e.g., Barrouillet & Lépine, 2005; Davis & Carr, 2002; Geary, 1994; Geary & Brown, 1991; Geary, Brown, & Samaranayake, 1991; Geary et al., 1999; Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Mabbott & Bisanz, 2003; Noël, Seron, & Trovarelli, 2004; Siegler, 1987, 1996; Steel & Funnell, 2001; Svenson & Sjöberg, 1983), including direct memory retrieval (e.g., 'knowing' that $8 + 5 = 13$), transformation (e.g., $8 + 6 = 8 + 2 + 4 = 10 + 4 = 14$), and counting (e.g., $4 + 3 = 4$… 5… 6… 7).

Unfortunately, all studies mentioned included a choice condition only, that is, a condition in which the children were free to choose any strategy they wanted. It has been shown convincingly that choice conditions provide reliable measures of strategy selection but not of strategy efficiency (Siegler & Lemaire, 1997). Indeed, strategy efficiency measures are biased by the strategy selection process. Because the current study aimed to investigate the role of working memory in both strategy selection and strategy efficiency, the choice/no-choice method (devised by Siegler & Lemaire, 1997) was used. This method includes a choice condition plus several no-choice conditions, in which participants are asked to use one single strategy for all problems. Data obtained in no-choice conditions provide reliable strategy efficiency measures. Some recent studies applied the choice/no-choice method successfully to investigate children's arithmetic performance (e.g., Carr & Davis, 2001; Lemaire & Lecacheur, 2002; Torbeyns, Verschaffel, & Ghesquière, 2002, 2004a, 2005a).

A third and final shortcoming is that very few studies investigated the role of working memory in normally achieving children (but see Adams & Hitch, 1997; Ashcraft & Fierman, 1982; Bull & Scerif, 2001; Geary, Bow-Thomas, Liu, & Siegler, 1996a; Hecht, Torgesen, Wagner, & Rashotte, 2001; Kaye et al., 1989). Because we believe it is important to know how the interaction between working memory and arithmetic performance progresses in normal development, the current study tested children without mathematical disabilities. A similar research question was raised by Barrouillet and Lépine (2005), who tested normally developing elementary-school children. They observed that children with high working memory capacities solved simple-addition problems more efficiently than did children with low working-memory capacities. Working-memory capacity correlated with strategy selection as well; percentages retrieval use were higher in high-capacity children than in low-capacity children.

To summarize, the current study addressed the development of working-memory involvement in children's arithmetic strategy use. To this end, an experiment combining the dual-task method and the choice/no-choice method was administered to 10-, 11-, and 12-year-old children. The dual-task method permits an online investigation of working-memory involvement in arithmetic performance, and the choice/no-choice method permits collection of reliable strategy selection and strategy efficiency data. These methods have been combined successfully in adult studies (Imbo & Vandierendonck, in press b,c) but not yet in child studies. However, results obtained in adult studies cannot simply be generalized to children. Therefore, the current study not only investigated the development of working-memory involvement in children's strategies but also tested whether or not results obtained in adult studies apply to children.

Our hypotheses are based on the assertion that many working-memory resources are needed during the initial phase of learning and that fewer working-memory resources are needed as procedural strategies (transformation and counting) are used less frequently and arithmetic facts become represented in long-term memory (see also Ackerman, 1988, Geary et al., 2004; Siegler, 1996). We suppose, however, that the decrease of working-memory involvement in arithmetic tasks across development is not caused by strategy selection processes only but rather is also caused by strategy efficiency processes.

First, age-related differences in *strategy selection* might change the ratio between working-memory involvement and the demands of the arithmetic task. Because direct memory retrieval needs fewer working-memory resources than do nonretrieval strategies, more frequent retrieval use might reduce the requirements of the arithmetic task, leaving more working-memory resources free for the secondary task. Stated differently, the impact of a working-memory load on the arithmetic task will diminish when strategy selection becomes more efficient (i.e., when the outcome of the selection process leads to the least demanding strategy).

Second, the ratio between working-memory involvement and simple-arithmetic task demands might be changed further by more efficient *retrieval* use. Because direct memory retrieval relies on working-memory resources (Imbo & Vandierendonck, in press b,c), it is hypothesized that faster retrieval would need fewer working-memory resources than would slow and effortful retrieval. Indeed, as problem-answer associations become stronger across development, fewer working-memory resources would be needed to retrieve the correct solution from long-term memory.

Third, we hypothesized that an age-related increase in *nonretrieval* strategy efficiency would also change working-memory involvement. Because nonretrieval strategies (transformation and counting) rely heavily on working-memory resources (Imbo & Vandierendonck, in press b,c), it is hypothesized that more efficient procedural use would need fewer working-memory resources than would less efficient procedural use. The componential steps used in nonretrieval strategies would become more practiced and require less effort with age, resulting in lower working-memory demands. The latter two hypotheses imply an age-related decrease in the impact of working-memory load on strategy efficiency. More specifically, we anticipate that the execution time of retrieval, transformation, and counting strategies will suffer less from a working-memory load as children become older.

Finally, we expected an age-related decrease in the working-memory costs due to general (i.e., non-mathematical) processes such as encoding stimuli and pronouncing answers. To test this prediction, a 'naming' condition was included in the current study. In this condition, children needed to name the correct answer to the problem presented on the screen. It was expected that the naming task would require fewer working-memory resources with growing age. The naming condition also offers the opportunity to test whether direct memory retrieval relies on working memory. If the impact of working-memory load on retrieval is larger than on naming, one may conclude that the very specific fact retrieval processes (i.e., long-term memory access, activation of the correct answer, and inhibition of incorrect answers) need working-memory resources.

INDIVIDUAL DIFFERENCES IN CHILDREN'S ARITHMETIC STRATEGY USE

To enhance understanding of children's arithmetic strategy use, the current study examined individual differences as well. Five individual-difference variables that might influence children's arithmetic performance were selected: short-term memory, processing speed, arithmetic skill, math anxiety, and gender.

Short-term memory. Short-term memory is a system that passively stores information and can be distinguished from working memory (which entails both storage and processing) already from 7 years of age on (Kail & Hall, 2001). Although the relation between short-term memory and arithmetic ability in mathematically disabled children is still questioned, short-term memory is not expected to play a great role in normally achieving children's arithmetic ability. Bull and Johnston (1997), for example, observed no correlations between short-term memory and retrieval frequency, retrieval efficiency, or counting efficiency. In the present study, the digit span test was used in order to collect data on children's short-term capacity.

Processing speed. The relation between processing speed and arithmetic ability was first examined by Bull and Johnston (1997). These authors observed that processing speed was – among several other variables such as short-term memory, speech rate, and item identification – the best predictor of mathematical ability. This result was further confirmed by Kail and Hall (1999), who observed that processing speed had the strongest and most consistent relation to arithmetic problem solving. Hitch, Towse, and Hutton (2001), in contrast, maintain that working-memory span is a better predictor of arithmetic ability than processing speed. In a longitudinal study by Noël et al. (2004), processing speed did not predict children's later performance on addition tasks. However, the researchers observed a bizarre correlation between processing speed and retrieval frequency in that slower participants were those who used retrieval more frequently. Thus, the evidence is equivocal concerning the role of processing speed as a critical determinant of simple-arithmetic performance. Because efficient strategy execution is generally defined as fast (and correct) strategy execution, we expected a positive correlation between processing speed and strategy efficiency. Because efficiently executed strategies strengthen the problem-answer association in long-term memory, we further expected that children with a higher processing speed would use retrieval more frequently. This expectation is in disagreement with the observation of Noël et al. (2004) but is more compelling than expecting a negative correlation between processing speed and retrieval frequency.

Arithmetic skill. The relation between arithmetic skill, on the one hand, and strategy selection and strategy efficiency, on the other, is straightforward in that persons who use retrieval frequently and who are fast in executing strategies will perform better on arithmetic skill tests. This relation has been shown in adults (e.g., Ashcraft, Donley, Halas, & Vakali, 1992; Campbell & Xue, 2001; Hecht, 1999; Imbo, Vandierendonck, & Rosseel, in press f; Kirk & Ashcraft, 2001; LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b) as well as in children (e.g., Geary & Burlingham-Dubree, 1989). We expected more frequent retrieval use and more efficient strategy use in high-skill children than in low-skill children.

Math anxiety. In adults, math anxiety is an individual-difference variable that affects online performance in math-related tasks (Ashcraft & Kirk, 2001). High and low-anxious adults differ in complex-arithmetic tasks (e.g., sums of two 2-digit numbers) but not in simple-arithmetic tasks (Ashcraft, 1995; Ashcraft & Faust, 1994; Faust, Ashcraft, & Fleck, 1996). More recently, however, effects of math anxiety have been observed on simple-arithmetic strategy use in adults (Imbo & Vandierendonck, in press c). In general, high-anxious adults were slower in the execution of both retrieval and nonretrieval strategies. Effects of math anxiety on strategy selection were also found in that percentages retrieval use were lower in high-anxious adults than in low-anxious adults. In the current child sample, high-anxious children were expected to be less efficient than low-anxious children, and high-anxious children were expected to use retrieval less often than low-anxious children.

Gender. Several studies indicated that gender differences exist in arithmetic strategy choices made by elementary-school children. More specifically, direct memory retrieval is chosen more frequently by boys whereas nonretrieval strategies are chosen more frequently by girls (Carr, 1996; Carr & Jessup, 1997; Davis & Carr, 2002). With respect to strategy efficiency, gender differences exist as well in that boys are faster than girls in executing computational processes (Carr & Jessup, 1997; Carr, Jessup, & Fuller, 1999; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Geary, Bow-Thomas, Fan, & Siegler, 1993a; Geary, Hamson, & Hoard, 2000a) and, more specifically, in direct memory retrieval (Royer, Tronsky, Chan, Jackson, & Marchant, 1999a). Based on these previous results, we expected more frequent and more efficient retrieval use in boys than in girls.

METHOD

PARTICIPANTS

Sixty-three children participated in the present study. They all attended the same elementary school in the Flemish part of Belgium. Twenty-one of them were in the 4th grade of elementary school (mean age: 10 years 0 months, 9 girls and 12 boys), twenty-one other children were in the 5th grade of elementary school (mean age: 11 years 1 month, 10 girls and 11 boys), and the last twenty-one children were in the 6th grade of elementary school (mean age: 12 years 2 months, 14 girls and 7 boys). Children were selected from the whole ability range, although those who were considered by their teachers to have specific learning or behavioral difficulties were excluded. The children had no documented brain injury or behavioral problems. The children only participated when they, as well as their teachers and their parents, consented.

PROCEDURE

Several individual-difference tests and one dual-task experiment were administered to each child. The whole procedure (individual-difference tests and dual-task experiment) took approximately one hour per child but was divided into two parts of 30 minutes each. Each child was tested individually in a quiet room. Testing started with short questions about the child such as age, grade (4^{th}, 5^{th}, or 6^{th}), and math anxiety (on a rating scale from 1 "low" to 5 "high"). Then, the first part of the dual-task experiment was run, after which the digit span test was administered. Approximately five days later, the second part of the dual-task experiment was run, after which the processing speed test was administered. After all individual experiments were run, the arithmetic skill test was run classically. Each individual-difference test and the dual-task experiment (consisting of a primary task and a secondary task) are described more extensively in the remainder of this section.

Primary task: solving simple addition problems. Children needed to solve simple addition problems in five conditions: a choice condition, three no-choice conditions (the order of which was randomized) and a naming condition (in which correct answers were presented on the screen). The choice condition always was the first so as to exclude influence of no-choice conditions on the choice condition, and the naming condition always was last so as to exclude effects of naming on solving the problems. In the choice condition, 6 practice problems and 32 experimental problems were presented. The no-choice conditions started immediately with the 32 experimental problems. Each condition was further divided in two blocks: a control block without working-memory load and a block in which the executive component of working memory was loaded. For half of the children, each condition started with the no-load block and was followed by the working-memory load block. The order was reversed for the other half of the children.

The addition problems were composed of pairs of numbers between 2 and 9, of which the sum exceeded 10 (e.g., 6 + 7). Problems involving 0 or 1 as an operand or answer (e.g., 5 + 0) and tie problems (e.g., 8 + 8) were excluded. Since commuted pairs (e.g., 9 + 4 and 4 + 9) were considered as two different problems, this resulted in 32 addition problems (ranging from 2 + 9 to 9 + 8). A trial started with a fixation point for 500 msec. Then the addition problem was presented horizontally in the center of the screen, with the plus sign (+) at the fixation point. In the naming condition, the problem was presented with its correct answer (e.g., "9 + 8 = 17"). The problem remained on screen until children responded. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, children wore a microphone that was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 msec. On each trial, feedback was presented to the children – a happy face when their answer was correct and a sad face when it was not.

Immediately after solving each problem, children in the choice condition were presented with four strategies on the screen (see e.g., Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler, Kirk, & Ashcraft, 2003): 'Retrieval', 'Count', 'Transform', and 'Other'. These four choices had been extensively explained by the experimenter. *Retrieval: You solve the problem by remembering or knowing the answer directly from memory. Count: You solve the problem by counting a certain number of times to get the answer. Transform: You solve the problem by referring to related operations or by deriving the answer from known facts. Other: You solve the problem by a strategy unlisted here, or you do not know what strategy that you used to solve the problem.* Examples of each strategy were presented as well. Children needed to report verbally which of these strategies they had used.

In the no-choice conditions, children were forced to use one particular strategy to solve all problems. In no-choice/retrieval, they were asked to retrieve the answer, in no-choice/transform, they were asked to transform the problem by making an intermediate step to 10 (e.g., 9 + 6 = 9 + 1 + 5 = 10 + 5 = 15), and in no-choice/count, they had to count (subvocally) until they reached the correct total (e.g., 7 + 4 = 7… 8… 9… 10… 11). Children were free to choose whether they started to count from the larger addend on (cf. the ‘min’ counting strategy, Groen & Parkman, 1972). After solving the problem, children also answered with ‘yes’ or ‘no’ to indicate whether they had succeeded in using the forced strategy. In choice and no-choice conditions, the children’s answer, the strategy information, and the validity of the trial were recorded online by the experimenter. All invalid trials (e.g., failures of the voice-activated relay) were discarded and returned to at the end of the block, thereby minimizing data loss due to unwanted failures.

Secondary task: executive working-memory load. An adapted version of the Continuous Choice Reaction Time Task – Random (CRT-R task; Szmalec, Vandierendonck, & Kemps, 2005) was used to load the executive working-memory component. Compared with the original version of the CRT-R task, the difference between low and high tones was larger (262 and 1048 Hz vs. 262 and 524 Hz), the interval between both tones was longer (2000 and 2500 msec vs. 900 and 1500 msec), and the duration of each tone was longer (300 msec vs. 200 msec). Children needed to press the 4 on the numerical keyboard when they heard a high tone and needed to press the 1 when they heard a low tone. This task was also performed alone (i.e., without the concurrent solving of addition problems) at the beginning of the working-memory load block.

Digit span. Digit span was tested using the Wechsler Intelligence Scale for Children-Revised (WISC-R) digit span subtest (Wechsler, 1986). In this task, digits are read aloud by the experimenter, and children need to repeat them in the correct order. There were two trials for each span length. The experimenter started from a span length of two digits and continued until the children made a mistake in both trials of the same span length. The highest span length reached by the children was set as 'digit span'.

Processing speed. Processing speed was tested by a visual number matching task (also used by Bull & Johnston, 1997[3]), which consisted of 30 rows of six digits, with two digits in each row being identical (for example 5 3 1 8 9 3). Children were instructed to cross out the identical digits in each row and to work both as quickly and as accurately as possible. The performance measure was the time taken to complete all 30 rows of digits. Note that a higher measure indicates a slower performance.

Arithmetic skill. A standardized skill test (Arithmetic Tempo Test; De Vos, 1992) was administered classically after all individual experiments were run. This pen-and-paper test consists of several subtests that require elementary computations. Each subtest concerns only one arithmetic operation. In the current experiment, we administered the addition subtest (e.g., 2 + 3 = ?; 76 + 18 = ?) and the subtraction subtest (7 – 5 = ?; 54 – 37 = ?), each consisting of 40 items of increasing difficulty. For each subtest, children were given 1 minute to solve as many problems as possible within that minute. Performance was the sum of the addition and the subtraction subtests.

RESULTS

Of all trials, 5.2% was spoiled due to failures of the sound-activated relay. Because these invalid trials were readministered at the end of the block, most of them were recovered, thereby reducing to 0.8% the trials spoiled due to failures of the sound-activated relay. Furthermore, all incorrect trials (3.5%), all choice trials on which children reported having

---

[3] We are grateful to these authors for providing us with the stimuli used in their visual number matching task.

used an Other strategy (0.3%), and all no-choice trials on which children failed to use the forced strategy (8.8%) were deleted. All data were analyzed on the basis of the multivariate general linear model, and all reported results are considered to be significant at $p < .05$ unless mentioned otherwise.

This section is divided into four parts. We start with the results of the secondary task. Thereafter, the results concerning strategy efficiency and strategy selection are reported. Finally, the importance of individual differences is discussed. Due to voice-key problems, two participants (one 4^{th} grader and one 6^{th} grader) were excluded from analyses, leaving scores for twenty 4^{th} graders, twenty-one 5^{th} graders, and twenty 6^{th} graders.

SECONDARY TASK PERFORMANCE

A 3 x 6 analysis of variance (ANOVA) was conducted on accuracy on the CRT-R task, with Grade (4^{th}, 5^{th}, 6^{th}) as between-subjects factor and Primary task (no primary task, naming, no-choice/retrieval, no-choice/transform, no-choice/count, choice) as within-subjects factors (see Table 1). The main effect of Grade was significant, $F(2,58) = 5.77$, $MS_e = 3770$; 4^{th} graders were less accurate than 5^{th} graders, $F(1,58) = 8.95$ but there was no difference between 5^{th} and 6^{th} graders, $F(1,58) < 1$. The main effect of Primary task was significant as well, $F(5,54) = 16.53$, $MS_e = 270$. Executing the CRT-R task without the primary task resulted in greater accuracy than CRT-R performance during naming, $F(1,58) = 4.49$, which in turn led to greater accuracy than CRT-R performance during no-choice/retrieval, $F(1,58) = 53.49$. Accuracy did not differ between no-choice/retrieval, no-choice/transform, and choice conditions, all $Fs(1,58) < 1$, but CRT-R accuracy was lower in the latter three conditions than in the no-choice/count condition, $F(1,58) = 10.47$, $F(1,58) = 4.01$, and $F(1,58) = 13.36$, respectively.

Table 1
Mean accuracies (%) and mean correct RTs (in msec) on the CRT-R task as a function of Grade and Primary task. Standard errors are shown between brackets.

Accuracies	4th grade	5th grade	6th grade
No primary task	46 (9)	76 (8)	77 (9)
Naming	42 (7)	66 (7)	70 (7)
No-choice/retrieval	26 (6)	43 (6)	46 (6)
No-choice/transform	29 (6)	48 (6)	45 (6)
No-choice/count	26 (4)	59 (6)	50 (6)
Choice	25 (6)	44 (5)	44 (6)
RTs	**4th grade**	**5th grade**	**6th grade**
No primary task	772 (49)	739 (48)	743 (49)
Naming	944 (54)	786 (52)	790 (54)
No-choice/retrieval	1052 (50)	1028 (48)	1066 (50)
No-choice/transform	1026 (47)	1024 (46)	1010 (47)
No-choice/count	980 (40)	975 (39)	1001 (40)
Choice	1034 (36)	1054 (35)	964 (36)

A similar 3 x 6 ANOVA was conducted on correct reaction times (RTs) of the CRT-R task (see Table 1). The main effect of Grade did not reach significance, $F(2,58) < 1$, $MS_e = 75586$, but the main effect of Primary task did, $F(5,54) = 25.99$, $MS_e = 36387$. Executing the CRT-R task without the primary task was faster than CRT-R performance during naming, $F(1,58) = 4.50$, which in turn was faster than performance during no-choice/retrieval, $F(1,58) = 24.78$. There were no significant differences in RTs among the no-choice/retrieval, no-choice/transform, no-choice/count, and choice conditions, all $Fs(1,58) < 1$, except that CRT-R performance was faster in no-choice/count than in no-choice/retrieval, $F(1,58) = 4.24$. The Grade x Primary task interaction was not significant, $F(10,110) < 1$.

STRATEGY EFFICIENCY

Because accuracy was very high, (100% in no-choice/naming, 97% in no-choice/retrieval, 98% in no-choice/transform, 98% in no-choice/count, and 95% in choice), strategy efficiency was analyzed in terms of strategy speed. Only the RTs uncontaminated by strategy choices (i.e., no-choice RTs) were considered. A 3 x 2 x 4 ANOVA was conducted on correct RTs with Grade (4^{th}, 5^{th}, 6^{th}) as between-subjects factor and Load (no load vs. load) and Task (naming, retrieval, transformation, counting) as within-subjects factors (see Table 2).

The main effect of Load was significant, $F(1,58) = 83.53$, $MS_e = 221390$, with higher RTs under load than under no-load. The main effect of Grade was also significant, $F(2,58) = 8.17$, $MS_e = 4145150$. Fourth graders were significantly slower than 5^{th} graders, $F(1,58) = 9.30$, but there was no difference between 5^{th} and 6^{th} graders, $F(1,58) < 1$. Finally, the main effect of Task was significant as well, $F(3,56) = 104.56$, $MS_e = 1451894$. Naming was faster than retrieval, $F(1,58) = 297.93$, retrieval was faster than transformation, $F(1,58) = 43.02$, and transformation was faster than counting, $F(1,58) = 28.62$.

Table 2
Mean corrects RTs (in msec) on the simple-arithmetic task (in no-choice conditions) as a function of Grade, Load, and Task. Standard errors are shown between brackets.

		4th grade	5th grade	6th grade
Naming	No load	641 (22)	525 (22)	511 (22)
	Load	989 (52)	794 (50)	790 (52)
Retrieval	No load	1650 (88)	1382 (86)	1115 (88)
	Load	2122 (106)	1763 (103)	1387 (106)
Transformation	No load	2550 (241)	2006 (235)	1816 (241)
	Load	3177 (325)	2610 (317)	2404 (325)
Counting	No load	4299 (360)	2684 (351)	2561 (360)
	Load	4778 (376)	2955 (367)	2644 (376)

Task further interacted with Grade, $F(6,114) = 4.08$ and with Load, $F(3,56) = 3.68$. The Task x Grade interaction indicated that the decrease in RTs over grades differed across strategies. Naming RTs decreased from 4th to 5th grade, $F(1,58) = 12.08$ but did not change from 5th to 6th grade, $F(1,58) < 1$. Retrieval RTs, in contrast, decreased from 4th to 5th grade, $F(1,58) = 5.95$ and from 5th to 6th grade, $F(1,58) = 6.24$. Transformation RTs did not change from 4th to 5th grade, $F(1,58) = 2.10$ or from 5th to 6th grade, $F(1,58) < 1$. Finally, counting RTs decreased from 4th to 5th grade, $F(1,58) = 11.86$ but not from 5th to 6th grade, $F(1,58) < 1$.

The Task x Load interaction showed that the effect of working-memory load (i.e., RT load – RT no-load) was the largest on transformation

RTs (606 msec). This effect was larger than the effects on naming RTs (299 msec), $F(1,58) = 10.58$, retrieval RTs (375 msec), $F(1,58) = 5.39$, and counting RTs (278 msec), $F(1,58) = 7.57$. As hypothesized, the effect of load was larger on retrieval RTs than on naming RTs, $t(58) = 1.87$, indicating that the retrieval process requires extra executive working-memory resources. It should be noted, however, that the effect of load was significant in each single task; $F(1,58) = 122.59$ for naming, $F(1,58) = 106.45$ for retrieval, $F(1,58) = 43.73$ for transformation, $F(1,58) = 7.28$ for counting.

The Grade x Load and Grade x Load x Task interactions did not reach significance, $F(2,58) = 1.40$ and $F(6,114) < 1$, respectively. Planned comparisons were conducted, however, to test the development of working-memory involvement in the different strategies. Whereas the effect of load on naming RTs did not change linearly[4] across grades, $F(1,58) = 1.08$, the effect of load on retrieval RTs decreased linearly across grades, $F(1,58) = 4.91$, with load effects of 472 msec, 382 msec, and 273 msec for 4th, 5th, and 6th grade, respectively. The effect of load on transformation RTs did not change either, $F(1,58) < 1$. Finally, the effect of load on the counting strategy tended to decrease linearly, $t(58) = 1.56$, $p = .062$, one-tailed, with load effects of 479 msec, 270 msec, and 83 msec, for 4th, 5th, and 6th grade, respectively.

To summarize, children require executive working-memory resources to solve simple addition problems. Even the simple task of saying an answer displayed on the screen ("naming") relies on executive resources. Retrieving an answer from long-term memory, however, needs even more executive resources. As children grow older, they become more efficient (faster) in the execution of retrieval and counting strategies but not in the execution of the transformation strategy. Increases in strategy efficiency are

---

[4] To test whether RTs did change linearly across grades, contrast values were -1 for 4th grade, 0 for 5th grade, and 1 for 6th grade.

accompanied by decreases in working-memory involvement. More specifically, higher retrieval and counting efficiencies reduced the requirements of executive resources, so that the negative impact of an executive load decreased with age. The executive resources needed in the naming task, however, remained the same across grades. The role of working memory in the transformation strategy (which relied most heavily on executive resources) did not change across grades either; all children relied equally heavily on their working memory to use this strategy.

STRATEGY SELECTION

To investigate effects on strategy selection, a 3 x 2 x 3 ANOVA was conducted on percentages strategy use (in the choice condition), with Grade (4^{th}, 5^{th}, 6^{th}) as between-subjects factor and Load (no load vs. load) and Strategy (retrieval, counting, and transformation) as within-subjects factors (see Table 3).

Table 3
Mean percentages strategy use (in the choice condition) as a function of Grade and Load. Standard errors are shown between brackets.

		4^{th} grade	5^{th} grade	6^{th} grade
Retrieval	No load	46 (7)	67 (7)	60 (7)
	Load	48 (7)	76 (7)	60 (7)
Transformation	No load	41 (6)	15 (6)	22 (6)
	Load	40 (6)	6 (6)	22 (6)
Counting	No load	12 (5)	18 (5)	18 (5)
	Load	12 (5)	15 (4)	18 (5)

The main effect of Strategy was significant, $F(2,57) = 31.91$, $MS_e = 2059$. Retrieval was used more frequently than transformation, $F(1,58) = 25.28$, which in turn was used more frequently than counting, $F(1,58) = 3.70$. Strategy further interacted with Grade, $F(4,116) = 2.64$. Retrieval use increased between 4^{th} and 5^{th} grade, $F(1,58) = 6.85$ but did not change between 5^{th} and 6^{th} grade, $F(1,58) = 1.63$. Transformation use decreased between 4^{th} and 5^{th} grade, $F(1,58) = 10.79$ but did not change between 5^{th} and 6^{th} grade, $F(1,58) = 1.31$. Finally, counting was used equally often between 4^{th} and 5^{th} grades and between 5^{th} and 6^{th} grades, both $Fs(1,58) < 1$. The Load x Strategy and Load x Strategy x Grade interactions did not reach significance.

To summarize, all strategies were used by the children, although retrieval was used more frequently than were transformation and counting. Retrieval use also increased as children grew older. No effects of load on strategy selection were observed.

INDIVIDUAL DIFFERENCES

Table 4 displays means of each individual-difference variable for each grade. The results of a one-way ANOVA, with Grade as between-subjects variable, are displayed in this table as well. The main effect of Grade was significant for arithmetic skill and processing speed but not for digit span or math anxiety. Planned comparisons showed that the progress in arithmetic skill and processing speed was significant between 4^{th} and 5^{th} grade but not between 5^{th} and 6^{th} grade.

To test the influence of individual differences on children's arithmetic strategy use, correlations between strategy efficiencies, strategy selection and the individual differences were calculated (see Table 5). To consolidate the results presented in the previous sections, working-memory load was also included in these correlational analyses. Gender was coded as a dummy variable; girls were coded as -1 and boys were coded as 1. Grade was coded as two dummy variables. For the first one (4th vs. 5th grade), 4th graders were coded as -1, 5th graders were coded as 1, and 6th graders were coded as 0. For the second one (5th vs. 6th grade), 4th graders were coded as 0, 5th graders were coded as -1, and 6th graders were coded as 1. Working-memory load was coded as a dummy variable as well: no-load was coded as -1 and load was coded as 1.

Table 4
Means and standard deviations of the individual-difference variables across grades. Results of the ANOVAs with Grade as between-subjects factor are displayed in the three rightmost columns.

	4th grade mean (SD)	5th grade mean (SD)	6th grade mean (SD)	$F(2,58)$ main effect	$F(1,58)$ 4th vs. 5th	$F(1,58)$ 5th vs. 6th
Digit span	5.5 (1)	5.8 (1)	5.7 (1)	0.66	1.28	0.04
Processing speed	96 (14)	79 (15)	72 (10)	16.80**	15.73**	3.08°
Arithmetic skill	46 (5)	53 (7)	57 (6)	19.76**	15.99**	2.84°
Math anxiety	2.3 (1)	2.4 (1)	2.1 (1)	0.42	0.15	0.85

° $p < .10$ * $p < .05$ ** $p < .01$

Table 5
Correlations between naming RTs, retrieval RTs, transformation RTs, counting RTs, percentages retrieval use, working-memory load, and the individual-difference variables. Note that efficiency and speed measures are expressed in RTs, so higher RTs indicate lower efficiencies.

	Retrieval Eff.	Transf. Eff	Count Eff.	Retrieval %	Digit Span	Proces. Speed	Arithm. Skill	Math Anxiety	Gender	4^{th} vs. 5^{th} grade	5^{th} vs. 6^{th} grade	WM Load
Naming Eff.	.52**	.42**	.37**	-.08	.00	.34**	-.30**	-.12	.12	-.26**	-.01	.62**
Retrieval Eff.		.62**	.61**	.00	-.01	.39**	-.59**	-.05	.11	-.24**	-.24**	.35**
Transf. Eff.			.61**	-.11	.01	.35**	-.56**	.02	.20*	-.17	-.06	.23*
Count Eff.				-.07	-.02	.39**	-.49**	.03	.02	-.38**	-.04	.08
Retrieval %					-.02	-.27**	.32**	-.28**	.18*	.32**	-.16	.05
Digit Span						-.23*	.14	-.20*	-.11	.15	-.05	---
Proces. Speed							-.58**	-.07	.30**	.42**	.18*	---
Arithm. Skill								-.19*	-.03	.05	-.12	---
Math Anxiety									-.25**	-.28*	.05	---
Gender										---	---	---
4^{th} vs. 5^{th} grade											---	---
5^{th} vs. 6^{th} grade												---

* $p < .05$ ** $p < .01$

The highest correlations appeared between the different types of strategy efficiency on the simple-arithmetic task (range .61 - .62). Children who efficiently retrieved simple-arithmetic facts from memory were also more efficient in employing nonretrieval strategies (counting and transformation).

Retrieval, transformation, and counting *efficiencies* further correlated with processing speed and arithmetic skill. Gender correlated with transformation efficiency only; transformation RTs were higher for boys than for girls. Fourth-grade children were slower than 5th grade children on naming, retrieval, and counting, but not on transformation. Fifth-graders were slower than 6th graders on the retrieval strategy only. Working-memory load correlated with naming RTs, retrieval RTs, and transformation RTs.

Strategy *selection* was also influenced by individual-difference variables. The retrieval strategy was used more frequently by children with higher processing speeds and higher arithmetic skills. Direct fact retrieval was used more frequently by 5th graders than by 4th graders, but it did not correlate with the contrast between 5th graders and 6th graders. Finally, retrieval use was higher in low-anxious children than in high-anxious children and was higher in boys than in girls.

Thus, the relations between strategy efficiency and strategy selection, on the one hand, and grade and working-memory load, on the other, are in agreement with the results reported previously. Children become more efficient in the execution of naming, retrieval, and counting strategies, whereas the efficiency of the transformation strategy does not increase across grades. The frequency of retrieval use also increases as children grow older. Finally, working-memory load predicted all strategy efficiencies except counting and did not predict strategy selection.

Table 5 revealed other noteworthy correlations as well. Math anxiety, for example, correlated with digit span and arithmetic skill; high-anxious children had lower digit spans and lower arithmetic skill scores. The correlation between math anxiety and digit span is in agreement with results obtained by Ashcraft and Kirk (2001), who observed that adults' working-memory span was negatively correlated with math anxiety. Although working memory cannot be equated with short-term memory, both results indicate that higher math anxiety scores go hand in hand with lower capacities for information storage and/or processing. Math-anxious participants are often occupied by worries and intrusive thoughts when performing arithmetic tasks (Ashcraft & Kirk, 2001; Faust et al., 1996). Because such intrusive thoughts load on storage and processing resources, high-anxious participants exhibit lower short-term-memory and working-memory capacities. The correlation between math anxiety and arithmetic skill corroborates the results obtained by Ashcraft (1995; Ashcraft & Faust, 1994; Ashcraft & Kirk, 2001; Faust et al., 1996), who observed that complex-arithmetic performance was worse in high-anxious adults than in low-anxious adults.

Gender correlated with processing speed and math anxiety. Girls scored higher on the math anxiety questionnaire than did boys. Girls were also faster on the processing speed task than were boys. The correlation between math anxiety and gender has been found previously; Ashcraft (1995) observed that highly anxious women (top quartile on anxiousness scale) scored almost one SD higher on a math-anxiety scale than highly anxious men. However, based on questionnaire results, it is impossible to rule out the possibility that females are just more honest in reporting their feelings than are males. The fact that girls were better on the processing speed test is in agreement with previous findings showing an advantage of females over males in perceptual speed (e.g., Kimura, 1992).

Subsequent hierarchical regression analyses assessed which variables contributed unique variance to the dependent variables naming efficiency, retrieval efficiency, transformation efficiency, counting efficiency, and retrieval frequency (see Appendix 1). In Model 1, we investigated whether the relation between the independent variables (arithmetic skill, working-memory load, processing speed, math anxiety, gender, and digit span) and the respective dependent variable was maintained when accounting for age-related changes (*df* = 1,118). Age-related changes indeed explained a large part of the variance; grade accounted for 10% of the variance in naming efficiency, $F(2,119) = 6.40$; for 24% of the variance in retrieval efficiency, $F(2,119) = 18.10$; for 6% of the variance in transformation efficiency, $F(2,119) = 3.62$; for 22% of the variance in counting efficiency, $F(2,119) = 16.94$; and for 10% of the variance in retrieval frequency, $F(2,119) = 6.55$.

In Model 1, we see that unique variance was found for arithmetic skill in predicting all four measures of simple-arithmetic strategic performance. Therefore, in Model 2 we investigated which variables were significant predictors when controlling for grade and arithmetic skill (*df* = 1,116). In Model 3 (*df* = 1,115), working-memory load was added to Model 2, whereas in Model 4 (*df* = 1,115), processing speed was added to Model 2.

Model 4 revealed that *working-memory load* contributed unique variance to naming efficiency, retrieval efficiency, and transformation efficiency, even when controlling for grade, arithmetic skill, and processing speed. However, working-memory load did not contribute unique variance to counting efficiency or retrieval frequency. *Processing speed* contributed unique variance to naming efficiency, transformation efficiency, and retrieval frequency when controlling for grade (Model 1). However, when working-memory load was entered into the model as well, processing speed was significant for naming efficiency only (Model 3). *Math anxiety* predicted retrieval efficiency and retrieval frequency. This contribution was significant even in Models 3 and 4. Finally, *gender* contributed unique variance to transformation efficiency (with boys being less efficient than girls), but this effect disappeared when controlling for processing speed (Model 4). However, gender did contribute unique variance to retrieval frequency in all four models.

Several results obtained in the hierarchical regression results stand out. First, although *processing speed* correlated with all four measures of simple-arithmetic strategic performance, processing speed did not contribute unique variance to any of these variables once working-memory load was entered into the analysis. Processing speed was significant for naming efficiency only. Second, *arithmetic skill* still contributed unique variance to the four simple-arithmetic performance measures when controlling for grade-related differences. However, arithmetic skill did not predict naming efficiency although both variables did correlate with each other. Third, partialing grade, arithmetic skill, and processing speed did not eliminate the significant role that *working memory* plays in predicting naming efficiency, retrieval efficiency, and transformation efficiency. Fourth, although *math anxiety* correlated with retrieval frequency only, regression analyses showed that it predicted both retrieval frequency and retrieval efficiency even in models 3 and 4. When free to choose the strategy they want (i.e., in choice conditions), high-anxious children used the retrieval strategy less frequently than did low-anxious children, but when high-anxious children were required to use retrieval (i.e., in no-choice/retrieval conditions) they sped up their retrieval use. Finally, the regression analyses uncovered a possible underlying cause of the correlation between *gender* and transformation efficiency. Given that girls had higher levels of processing speed than did boys (Table 5), the correlation between gender and transformation efficiency might be caused by gender differences in processing speed. Indeed, gender did not contribute unique variance to transformation efficiency when processing speed was entered into the analysis. However, gender contributed unique variance to retrieval frequency even in models 3 and 4. Retrieval use was more frequent in boys than in girls and this effect persisted even when controlling for grade, arithmetic skill, processing speed, and working-memory load.

GENERAL DISCUSSION

THE ROLE OF WORKING MEMORY IN CHILDREN'S STRATEGY EFFICIENCY AND STRATEGY SELECTION

The current results show that school-age children rely on working-memory resources to perform simple-arithmetic problems. Taxing children's executive working-memory resources resulted in poorer arithmetic performance; children of all ages executed strategies less efficiently. The impact of an executive working-memory load on children's *retrieval* efficiency is in agreement with comparable results obtained in adults (e.g., Anderson, Reder, & Lebiere, 1996; Imbo & Vandierendonck, in press b,c) and indicates that working-memory resources are needed to select information from long-term memory (Barrouillet & Lépine, 2005; Barrouillet, Bernardin, & Camos, 2004; Cowan, 1995, 1999; Lovett, Reder, & Lebiere, 1999). It is important to note that the impact of the executive working-memory load was larger when answers needed to be retrieved from long-term memory than when answers were provided (i.e., the 'naming' condition). Presumably, except for retrieval of the correct answer, the processes of digit encoding and pronouncing were equal in the naming condition and the retrieval condition. This result shows that retrieval of the correct answer and inhibition of incorrect answers do rely on executive working-memory resources. Recently, the executive function of inhibitory control has been shown to contribute to emergent arithmetic skills in preschool children (Espy et al., 2004).

The role of working memory was larger in *nonretrieval* strategies than in direct memory retrieval, a result obtained in adult studies as well (Imbo & Vandierendonck, in press b,c). Indeed, in addition to the fact that procedural strategies (transformation and counting) are composed of multiple retrievals from long-term memory, these strategies also contain several processes that might require *extra* executive resources such as performing calculations, manipulating interim results, and monitoring counting processes.

The arithmetic performances of normally developing children under executive working-memory load can be compared with arithmetic performances of mathematically impaired children, who are slower in solving arithmetic problems (e.g., Geary, 1993; Hitch & McAuley, 1991; Passolunghi & Siegel, 2001; Siegel & Linder, 1984; Siegel & Ryan, 1989; Swanson, 1993). This impairment has often been attributed to limitations in working memory and especially to limitations in the executive working-memory component (e.g., McLean & Hitch, 1999; Passolunghi & Siegel, 2004). That lower arithmetic performance can be caused by limitations in working memory was confirmed by the present results, in which executive working-memory resources (of normally developing children) were limited experimentally.

DEVELOPMENT OF THE ROLE OF WORKING MEMORY

The main goal of the current study was to investigate age-related changes in the ratio of available working-memory resources against simple-arithmetic task demands. The main conclusion is that the negative impact of an executive working-memory load decreases as children grow older. This conclusion corroborates the assertion that more working-memory resources are needed during the initial phases of skill acquisition and that fewer working-memory resources are needed with learning, namely when procedural strategies are used less frequently and retrieval strategies more frequently (Ackerman, 1988; Ackerman & Cianciolo, 2000; Geary et al., 2004; Siegler, 1996). Based on the results obtained in the current study, we infer that the declining impact of working-memory load is caused by age-related changes in strategy efficiency and strategy selection, but not by age-related changes in overall processing costs. These effects are discussed in the remainder of this section.

First, frequency of retrieval use increased across grades; 5th and 6th graders used retrieval more often than did 4th graders. Because direct memory retrieval is less effortful and requires fewer working-memory resources than do nonretrieval strategies such as counting and transformation (cf. no-choice data), more frequent retrieval use goes hand in hand with lower working-memory involvement. Phrased differently, more frequent retrieval use leaves more working-memory capacity free for other uses. This spare capacity can then be applied in the executive secondary task.

Second, retrieval efficiency increased across grades; direct memory retrieval took longer in 4th grade than in 5th grade and took longer in 5th grade than in 6th grade. More efficient retrieval use results from stronger problem-answer associations for the correct answer and weaker problem-answer associations for the neighboring incorrect answers. Stronger associations between the problem and its correct answer reduce the amount of executive working-memory resources needed to inhibit incorrect answers.

Third, counting efficiency increased across grades; counting was slower in 4th grade than in 5th and 6th grades. As counting becomes more efficient, fewer working-memory resources are needed, thereby reducing the working-memory involvement across ages. The increase in counting efficiency might be caused by increases in retrieval and procedural efficiency, increases in processing speed, and increases in speech rate. The faster children can count, the less information needs to be protected from decay. Importantly, transformation efficiency did not change across grades, and neither did the effect of working-memory load on transformation efficiency.

Finally, results showed that the age-related decline in the impact of working-memory load could not be due to developmental changes in overall processing costs. Although naming RTs were larger in 4th grade than in 5th and 6th grades, the effect of working-memory load on naming did not decrease with age. To conclude, the changing ratio between working-memory involvement, on the one hand, and simple-arithmetic performance, on the other, was due to age-related changes in strategy selection and strategy efficiency (for retrieval end counting), but not to age-related changes in general processes such as encoding and pronunciation.

Importantly, our conclusions are in agreement with a recent functional Magnetic Resonance Imaging (fMRI) study. Rivera, Reiss, Eckert, and Menon (2005) tested 8- to 19-year olds on arithmetic tasks and found that activation in the prefrontal cortex decreased with age, suggesting that younger participants need more working-memory and attentional resources to achieve similar levels of mental arithmetic performance. Activation of the hippocampus, the dorsal basal ganglia, and the parietal cortex decrease with age as well, suggesting greater demands on declarative, procedural, and visual memory systems in younger children than in older children (Qin et al., 2004; Rivera et al., 2005).

Future research may use the method adopted here (i.e., a combination of the dual-task method and the choice/no-choice method) to investigate *which* executive resources come into play in children's arithmetic strategy performance. Previous (correlational) research suggests that both inhibition and memory updating play a role in children's arithmetic problem solving (e.g., Passolunghi et al., 1999; Passolunghi & Pazzaglia, 2005).

INFLUENCE OF INDIVIDUAL DIFFERENCES ON CHILDREN'S STRATEGY USE

Digit span. Digit span did not correlate with strategy efficiency or strategy selection measures. This is at variance with previous studies in which a relation between short-term memory and arithmetic ability was observed (e.g., Hecht et al., 2001; Geary et al., 1991, 2000a; Hitch & McAuley, 1991; Passolunghi & Siegel, 2001; Siegel & Ryan, 1989; Swanson & Sachse-Lee, 2001). It should be noted, however, that a number of these studies included mathematically disabled children without taking reading ability or general intelligence into account. In as many other studies, no relation between short-term memory and mathematical ability was observed (e.g., Bayliss, Jarrold, Gunn, & Baddeley, 2003; Bull & Johnston, 1997; Geary et al., 2000a; Passolunghi et al., 1999; Passolunghi & Siegel, 2004; Swanson, 2006, Temple & Sherwood, 2002), in agreement with the current results. Individual differences in short-term memory apparently do not play an important role in children's simple-arithmetic performance. Individual differences in working memory, in contrast, do play a role in children's simple-arithmetic strategy use. Indeed, correlations between working-memory measures and mathematics ability have been found consistently (e.g., Bull et al., 1999; Bull & Scerif, 2001; Geary et al., 1999; McLean & Hitch, 1999; Noël et al., 2004; Passolungi et al., 1999; Passolunghi & Siegel, 2004; Swanson, 2004, 2006; Swanson & Beebe-Frankenberger, 2004; Swanson & Sachse-Lee, 2001, van der Sluis et al., 2004).

In our study with normally developing children, working memory (as loaded by the CRT-R task) but not short-term memory (as tested with the digit span) was related to arithmetic performance. Thus, we agree with Steel and Funnell (2001) in asserting that the number of items that can be stored in memory is less important than the ability to control attention and maintain information in an active, quickly retrievable state (see also Engle, 2002). The current results are also in agreement with most of the recent studies on mathematically disabled children. Children with arithmetic learning difficulties might suffer from a working-memory deficit (Geary, 2004) rather than a short-term memory deficit.

Processing speed. We observed that retrieval efficiency, transformation efficiency, counting efficiency, and retrieval frequency were lower in children with slower processing speed than in children with faster processing speed. Correlations between processing speed and arithmetic ability have been observed previously (e.g., Bull & Johnston, 1997; Durand, Hulme, Larkin, & Snowling, 2005; Kail & Hall, 1999). Kail and Hall (1999) hypothesized that faster processing is associated with faster retrieval of problem solving heuristics. The current research is consistent with their hypothesis in that we observed that faster processing children were more likely than slower processing children to select fast retrieval strategies. It should be noted, however, that the current results disagree with the results of Noël et al.'s (2004), who observed that slower processing children used retrieval more frequently than did faster processing children. Our results, however, are consistent with the expectation that faster processing children develop stronger problem-answer associations in long-term memory, resulting in more frequent retrieval use.

According to Bull and Johnston (1997), slow processing children may experience several difficulties. They may be slower in general information processing; however, they may also simply lack the automaticity to perform basic arithmetic operations. Based on the results obtained in the hierarchical regression analyses, the first explanation seems more plausible. Indeed, when controlling for age and working memory, processing speed did not contribute unique variance to any of the four arithmetic performance measures. Hence, the relation between processing speed and arithmetic performance is due to age-related speed and general working-memory deficits rather than to specific deficits in processing and automatizing numbers and number facts.

Arithmetic skill. High correlations between arithmetic skill, on the one hand, and strategy selection and strategy efficiency, on the other, were observed. Moreover, arithmetic skill contributed unique variance when partialing age from the analyses. Obviously, children who frequently use direct memory retrieval, retrieve answers from long-term memory efficiently, and execute nonretrieval strategies efficiently are in a good position to acquire general computational skills, resulting in good performance on general math attainment tasks. This agrees well with Hecht et al.'s (2001) finding that, in elementary-school children, simple-arithmetic efficiency is a significant predictor of later variability in general computational skills, even when controlling for phonological skills.

Math anxiety. Math anxiety did not correlate with the efficiency with which children used different strategies. This is in agreement with the assertion that math anxiety affects only complex-arithmetic performance and not simple-arithmetic performance (Ashcraft, 1995; Faust et al., 1996). Math anxiety did indeed correlate with performance on the (more complex) arithmetic-skill test; high-anxious children solved fewer problems than did low-anxious children, indicating more efficient complex problem solving in the latter than in the former.

Math anxiety further correlated with simple-arithmetic strategy selection; high-anxious children used retrieval less often than did low-anxious children. This effect of math anxiety on strategy selection can easily be explained on the basis of the strategy choice model of Siegler and Shrager (1984). In their model, each participant has his or her own confidence criterion. When solving simple-arithmetic problems, the strength of the problem-answer association is compared with this subjective confidence criterion. If the problem-answer associative strength exceeds the confidence criterion, the answer is emitted. If the problem-answer associative strength does not exceed the confidence criterion, then the child may continue to search memory for other candidate answers or may resort to a procedural strategy to compute the answer. If we suppose that anxious children set very high confidence criteria so as not to produce any incorrect answers, problem-answer associations will meet those criteria infrequently, resulting in less frequent retrieval use and more frequent procedural use.

Gender. Girls were more efficient in transformation use, whereas the retrieval strategy was used more frequently by boys. More frequent retrieval use in boys has been observed in previous studies (Carr, 1996; Carr & Jessup, 1997; Davis & Carr, 2002), and has been attributed to the effect of temperament (Davis & Carr, 2002). More efficient transformation use in girls than in boys has not been reported previously, but the current study showed that this observation might be related to gender differences in processing speed (cf. the hierarchical regression analyses). The more efficient transformation use in girls than in boys might also help to explain the gender difference in strategy selection; because girls are reasonably fast in applying the transformation strategy, they might opt not to switch to the retrieval strategy, which is only slightly faster for them. In boys, in contrast, the retrieval strategy is considerably faster than the transformation strategy, leading them to choose the fastest strategy (retrieval) more often. It is noteworthy that no gender differences were observed in retrieval efficiency. In previous studies, males were observed to be faster retrievers than females, both in children (Royer et al., 1999a) and adolescents (Imbo et al., in press f). Thus, more efficient retrieval use in boys than in girls is not found consistently across studies.

What causes such gender differences in arithmetic performance? According to Geary (1999) and Royer and colleagues (1999a), gender differences in arithmetic performance are not likely to be biologically based. Social and occupational interests seem to be a more reasonable cause. Royer and colleagues supposed that boys engage in out-of-school activities that provide them with additional practice on the manipulation of mathematical information. Geary, Saults, Liu, and Hoard (2000b) maintained that the male advantage in mathematical problem solving is due to a male advantage in spatial cognition. In sum, it is clear gender differences in arithmetic performance and their sources are not understood well and should be investigated further.

SUMMARY

In the present study, two approved methods were combined in order to investigate the development of working-memory involvement in children's arithmetic strategy use. The dual-task method permitted an online investigation of working-memory involvement in arithmetic performance, and the choice/no-choice method permitted achieving reliable strategy selection and strategy efficiency data. As far as we know, the combination of both methods has not yet been used in child studies. The results showed that, across development, the effect of an executive working-memory load decreased when retrieval was used more frequently and when strategies were executed more efficiently. However, the age-related decline in working-memory use was not due to developmental changes in other, more general processes, which required working-memory resources across all ages. Individual-difference variables (gender, math anxiety, arithmetic skill, and processing speed) accounted for differences in strategy selection and strategy efficiency as well. Arithmetic skill and working memory contributed more unique variance to arithmetic performance than processing speed and short-term memory did. Math anxiety and gender predicted some but not all of the arithmetic performance measures. Future research on working memory, strategy use, and mental arithmetic may investigate other arithmetic operations (subtraction, multiplication, division), other working-memory resources (phonological loop and visuo-spatial sketchpad), and other individual differences (e.g., motivation, intelligence, etc.)

Appendix 1
Hierarchical regression analyses for naming efficiency, retrieval efficiency, transformation efficiency, counting efficiency, and retrieval frequency.

	Model 1			Model 2			Model 3			Model 4		
Naming efficiency	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*
Arithm. skill	.020	2.68	-.181									
WM load	.382	86.63**	.618	.386	90.32**	.622				.385	94.41**	.621
Proces. speed	.041	5.66*	.256	.027	3.69	.219	.026	6.47*	.217			
Math anxiety	.017	2.22	-.130	.027	3.65	-.169	.026	6.34*	-.166	.018	2.44	-.140
Gender	.006	0.81	.081	.009	1.14	.096	.010	2.45	.105	.002	0.28	.049
Digit span	.002	0.20	.040	.003	0.35	.052	.003	0.68	.055	.007	0.92	.085

* $p < .05$ ** $p < .01$. Model 1 = age controlled (*df* per test= 1,118). Model 2 = age + arithmetic skill controlled (*df* per test= 1,116). Model 3 = age + arithmetic skill + working-memory load controlled (*df* per test= 1,115). Model 4 = age + arithmetic skill + processing speed controlled (*df* per test= 1,115).

Appendix 1 (continued)

	Model 1			Model 2			Model 3			Model 4		
Retrieval efficiency	*RΔ*	*F*	*Beta*	*RΔ*	*F*	*Beta*	*RΔ*	*F*	*Beta*	*RΔ*	*F*	*Beta*
Arithm. skill	.138	26.26**	-.474									
WM load	.124	22.85**	.353	.122	28.87**	.350				.122	28.62**	.350
Proces. speed	.016	2.58	.161	.000	0.00	.001	.000	0.00	.000			
Math anxiety	.007	1.03	-.082	.030	6.03	-.179	.030	7.41**	-.178	.032	6.24*	-.187
Gender	.000	0.01	.007	.001	0.17	.031	.001	0.30	.037	.001	0.19	0.34
Digit span	.001	0.15	.032	.004	0.68	.061	.004	0.89	.063	.004	0.70	0.63

* $p < .05$ ** $p < .01$. Model 1 = age controlled (*df* per test= 1,118). Model 2 = age + arithmetic skill controlled (*df* per test= 1,116). Model 3 = age + arithmetic skill + working-memory load controlled (*df* per test= 1,115). Model 4 = age + arithmetic skill + processing speed controlled (*df* per test= 1,115).

Appendix 1 (continued)

	Model 1			Model 2			Model 3			Model 4		
Transformation efficiency	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*
Arithm. skill	.208	35.24**	-.582									
WM load	.053	6.97**	.229	.048	8.58**	.218				.047	8.66**	.218
Proces. speed	.063	8.47**	.316	.014	2.42	.158	.014	2.55	.157			
Math anxiety	.000	0.01	.009	.018	3.15	-.139	.018	3.30	-.138	.013	2.22	-.119
Gender	.028	3.58	.171	.029	5.11	.177	.030	5.67*	.180	.020	3.52	.154
Digit span	.002	0.20	.040	.005	0.78	.069	.005	0.86	.070	.008	1.43	.094

* $p < .05$ ** $p < .01$. Model 1 = age controlled (*df* per test= 1,118). Model 2 = age + arithmetic skill controlled (*df* per test= 1,116). Model 3 = age + arithmetic skill + working-memory load controlled (*df* per test= 1,115). Model 4 = age + arithmetic skill + processing speed controlled (*df* per test= 1,115).

Appendix 1 (continued)

	Model 1			Model 2			Model 3			Model 4		
Counting efficiency	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*
Arithm. skill	.069	11.38**	-.335									
WM load	.006	0.88	.076	.006	1.01	.078				.006	1.00	.078
Proces. speed	.019	2.93	.173	.003	0.45	.070	.003	0.45	.069			
Math anxiety	.000	0.04	.016	.002	0.28	-.042	.002	0.27	-.042	.001	0.16	-.033
Gender	.003	0.45	-.056	.001	0.14	-.031	.001	0.13	-.030	.002	0.37	-.052
Digit span	.002	0.27	.043	.004	0.69	.065	.004	0.69	.066	.006	0.93	.078

* $p < .05$ ** $p < .01$. Model 1 = age controlled (*df* per test= 1,118). Model 2 = age + arithmetic skill controlled (*df* per test= 1,116). Model 3 = age + arithmetic skill + working-memory load controlled (*df* per test= 1,115). Model 4 = age + arithmetic skill + processing speed controlled (*df* per test= 1,115).

Appendix 1 (continued)

	Model 1			Model 2			Model 3			Model 4		
Retrieval frequency	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*	*R*Δ	*F*	*Beta*
Arithm. skill	.053	7.30**	.293									
WM load	.003	0.38	.054	.004	0.49	.060				.004	0.50	.060
Proces. speed	.030	4.13*	-.219	.010	1.43	-.135	.010	1.43	-.136			
Math anxiety	.091	13.31**	-.304	.065	9.58**	-.262	.064	9.51**	-.261	.079	11.96**	-.295
Gender	.038	5.15*	.199	.035	4.94*	.193	.035	4.97*	.194	.051	7.52**	.246
Digit span	.004	0.52	-.063	.006	0.89	-.082	.006	0.88	-.081	.010	1.41	-.104

* $p < .05$ ** $p < .01$. Model 1 = age controlled (*df* per test= 1,118). Model 2 = age + arithmetic skill controlled (*df* per test= 1,116). Model 3 = age + arithmetic skill + working-memory load controlled (*df* per test= 1,115). Model 4 = age + arithmetic skill + processing speed controlled (*df* per test= 1,115).

CHAPTER 8
GENERAL DISCUSSION

"Arithmetic is where the answer is right and everything is nice
and you can look out of the window and see the blue sky -
or the answer is wrong and you have to start all over and try again
and see how it comes out this time."

Carl Sandburg

When solving simple-arithmetic problems like 6 + 9 and 7 x 8, people use multiple strategies. Although direct memory retrieval is generally the most frequently used strategy, nonretrieval strategies are used as well. The importance of the current thesis lays in the fact that people use the available strategies neither equally frequently nor equally efficiently. How does this come? Which variables affect people's strategy use? In this final chapter, we try to answer these questions. First, the most important findings of the present thesis are summarized and put in a broader (brain-imaging) context. Next, we thoroughly discuss strategy efficiency and strategy selection. Then, the merits and limits of the current thesis are discussed. We also provide some practical implications of the current study. The chapter ends with some directions for future research.

RESEARCH OVERVIEW

PRACTICE EFFECTS

Practice effects were investigated in Chapters 2, 3, and 5. In Chapter 2, practice was investigated in a rather 'ecological' way: we tested secondary-school students with smaller and larger amounts of arithmetic lessons. The results showed that more frequent practice induced more frequent retrieval use, more efficient retrieval use, and more efficient nonretrieval use. Chapter 3 reported two experiments in which practice was experimentally manipulated. Here, we also observed practice effects on strategy efficiency and strategy selection. Retrieval was more frequently used after practice than before, and strategies were executed more efficiently after practice than before. Finally, based on the results obtained in Chapters 2 and 3, an experiential variable was included in Chapter 5. In this chapter, we observed more frequent retrieval use (Experiment 1) and more efficient strategy execution (Experiment 2) in participants with larger amounts of past arithmetic experiences.

Importantly, brain-imaging studies are in line with our results. In an Event Related brain Potential (ERP) study, Pauli et al. (1994) trained participants extensively on simple multiplications and observed that fronto-central positivity diminished over training sessions, indicating that participants tended to use a deliberate, conscious calculation strategy in the first sessions and direct fact retrieval in the last sessions. Delazer et al. (2003) trained participants extensively on complex-arithmetic problems. Functional Magnetic Resonance Imaging (fMRI) data showed a modification from quantity-based processing (cf. activation in the intraparietal sulcus) to more automatic retrieval (cf. activation in the angular gyrus), an observation that has been replicated more recently (e.g., Delazer et al., 2005; Ischebeck et al., 2006).

THE ROLE OF WORKING MEMORY

The role of working memory in simple-arithmetic strategy use was investigated by means of the selective interference paradigm in Chapters 4, 5, and 7, and by means of the individual-difference paradigm in Chapter 6. Executive working memory was needed for the efficient execution of both retrieval and nonretrieval strategies. Phonological working memory was only needed for the efficient execution of nonretrieval strategies. Neither executive nor phonological resources were needed in the strategy selection process.

The need for attentional resources in arithmetic strategy execution has also been confirmed by brain-imaging studies. In an fMRI study, Chochon, Cohen, van de Moortele, & Dehaene (1999) observed a strong activation in the prefrontal cortex during simple arithmetic, which they attributed to the use of nonretrieval strategies and the resulting involvement of working-memory resources. In a subsequent Positron Emission Tomography (PET) study, Zago et al. (2001) argued that frontal brain regions were more highly activated in nonretrieval strategies than in direct fact retrieval, which also confirms the involvement of executive working-memory resources during mental calculation (see also the fMRI study by Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001). Finally, Dehaene, Piazza, Pinel, and Cohen (2003) and Stanescu-Cosson et al. (2000) showed that direct memory retrieval activates left, parietal regions whereas nonretrieval strategies activate bilateral, frontoparietal regions.

DEVELOPMENTAL EFFECTS

In Chapters 6 and 7, we investigated strategy efficiency and strategy selection across development. Across age, retrieval use became more frequent while nonretrieval use vanished. The efficiency with which retrieval and nonretrieval strategies were executed also increased with age. These results are in agreement with the adaptive strategy choice model (ASCM, Siegler & Shipley, 1995), according to which children first switch from less efficient nonretrieval strategies to more efficient nonretrieval strategies, after which they switch from nonretrieval strategies to direct memory retrieval (i.e., through the construction and strengthening of problem-answer associations).

In the brain-imaging field, Kawashima et al. (2004) performed an fMRI study on children's and adults' addition, subtraction, and multiplication performance. Only very few brain areas showed significant activation differences between children and adults. Therefore, Kawashima et al. concluded that the cortical networks involved in simple arithmetic are similar between children and adults. In an fMRI study by Rivera, Reiss, Eckert, and Menon (2005), in contrast, significant age-related changes were observed. These researchers observed greater activation in the left parietal cortex and in the left occipital temporal cortex in older children and greater activation in the prefrontal cortex in younger children (see also Ansari, Garcia, Lucas, Hamon, & Dhital, 2005). This activation pattern suggests that young children require comparatively more working-memory and attentional resources to achieve similar levels of arithmetic performance – which is in agreement with our results.

STRATEGY EFFICIENCY

In this section, we discuss cross-experimental findings concerning people's strategy efficiency (i.e., the speed and accuracy with which retrieval and nonretrieval strategies are executed). The discussion is based upon several relevant questions that arose across this PhD project.

CAN PRACTICE EFFECTS BE MATHEMATICALLY DESCRIBED?

Practice enhanced the efficiencies with which strategies were executed. However, these increases were not equally large across strategies.

A frequently used law to describe practice effects on performance efficiency is the power-law of Newell and Rosenbloom (1981)[1]. This law predicts a negatively accelerating rate of speedup as a function of practice, and a larger speedup in early practice stages than in later practice stages. Based on our results, we argue that this law should not be applied to overall latencies, i.e., retrieval and nonretrieval latencies combined. In contrast, we believe that this law should be applied to retrieval and nonretrieval strategies separately. Delaney, Reder, Staszewski, and Ritter (1998) re-analyzed data obtained in previous practice studies with complex-arithmetic problems. They indeed observed that the improvement in latencies was better explained by practice *on a strategy* than by practice on the task as a whole. This has also been confirmed by Rickard (1997), who showed that the power law does not hold for practice effects on overall latency data, but does hold within each strategy (cf. the component power law or CMPL theory). In the current doctoral dissertation, larger practice effects in multiplication than in addition were observed. Hence, we argue that different parameters should be used when applying the power law to different operations.

CAN THE PROBLEM-SIZE EFFECT DISAPPEAR?

Although practice enhanced retrieval efficiency, the retrieval problem-size effect (i.e., retrieval latencies on large problems – retrieval latencies on small problems) never disappeared (see also Campbell & Xue, 2001; LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b; Penner-Wilger, Leth-Steensen, & LeFevre, 2002). Large problems must thus have some inherent difficulties in comparison with small problems. Candidates of such inherent differences are the fact that small problems have stronger problem-answer associations than do large problems, and the fact

---

[1] $T = BN^{-\alpha}$ in which T represents the performance time, B is the time taken to perform the first trial, N is the trial number, and α represents the rate at which performance time changes.

that small problems undergo less strong interference effects than do large problems (e.g., Ashcraft, 1992; Campbell, 1995; Geary, 1996; Zbrodoff, 1995). Can the problem-size effect then ever disappear? Yes, it can, albeit only in extraordinary cases. For example, Geary (1996) observed no problem-size effect in (probably extensively practiced) Chinese children, and Pauli, Lutzenberger, Birbaumer, Rickard, and Bourne (1996) observed no problem-size effect in a mental calculator. Answer magnitude thus produces inherent changes in strategy efficiencies, which can only be overcome, if at all, by very extensive practice.

WHEN DOES TRANSFER OCCUR?

In Chapter 3, we hypothesized that practice effects on nonretrieval strategy efficiency would transfer to other problems, whereas practice effects on retrieval efficiency would not. Although the results of our own experiments did not completely confirm this hypothesis (transfer occurred in Experiment 1 but not in Experiment 2), this hypothesis has been confirmed in various related studies. In an alphabet arithmetic task, Logan and Klapp (1991) observed that latencies on new, unpracticed problems were slower for participants who had learned by rote memory than for participants who had learned ‘by doing’. Comparably, Spelke and Tsivkin (2001) observed that practice generalized to new problems when using an approximation strategy but not when using an exact strategy. Delazer et al. (2005) also observed more accurate performance after practicing by nonretrieval strategies than after practicing by drill. Finally, Campbell, Fuchs-Lacelle, and Phenix (2006) observed that transfer between operations was eliminated when only retrieval trials were included.

The identical elements model (Rickard, 2005; Rickard & Bourne, 1996) states that transfer only occurs when both operands are identical. Hence, transfer occurs from 5 + 8 to 8 + 5 and from 3 x 7 to 7 x 3, but not from 13 – 8 to 13 – 5 or from 21 : 3 to 21 : 7. We agree with this model as developed for retrieval trials only. However, it might be interesting to extend the model in order to accommodate transfer effects associated with *non*retrieval strategies. Future adaptations of this model should thus differentiate between retrieval and nonretrieval transfer effects (see e.g., Rickard, 1997).

WHAT IS THE ROLE OF EXECUTIVE RESOURCES IN RETRIEVAL EFFICIENCY?

Before answering this question, one should ask what happens in our mind when we retrieve known facts from long-term memory. Let us return to the (non-arithmetic) example provided in the Introduction. When asked to retrieve the capital city of Papua-New-Guinea, people might directly retrieve the correct answer (Port Moresby) or an incorrect answer (e.g., Bissau). Otherwise, you might also have a "feeling of knowing" but fail to retrieve the answer quickly[2]. Anyhow, you will never (erroneously) say that Paris is the correct answer. Might there be similar processes active when trying to retrieve an arithmetic fact (e.g., 7 x 8) from long-term memory? Some people might directly know the correct answer (56); others will have a "feeling of knowing" and need some time to reach the correct answer. Still others will give an erroneous answer – that is mostly table-related (e.g., 64) whereas almost nobody will give the erroneous answer 14. This example enfeebles the notion that retrieval is always fast and correct. It is comparable with trying to recall the name of your first-grade teacher: people have to retrieve it from long-term memory but this retrieval might ask some time. In the following, we discuss the memory processes that take place when retrieving an answer from long-term memory. More specifically, we go more deeply into the question whether or not executive (or attentional) resources are needed in direct memory retrieval. Note that collecting strategy reports

[2] A *feeling of knowing* can be defined as "the degree of belief that a piece of information can be retrieved from memory" (Schunn, Reder, Nhouyvanisvong, Richards, & Stroffolino, 1997, p.4).

enabled us to differentiate between (possibly slow) retrieval processes and (possibly fast) nonretrieval strategies.

Whether or not direct fact retrieval relies on executive resources is a matter of debate. Some researchers believe that retrieval merely involves the obligatory activation of number facts and thus occurs automatically (i.e., without requiring cognitive resources; e.g., Ashcraft, 1987; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; LeFevre, Bisanz, & Mrkonjic, 1988; LeFevre & Kulak, 1994; Logan, 1985). Other researchers, in contrast, maintain that even the simplest fact retrieval requires attentional resources (e.g., Anderson, Reder, & Lebiere, 1996; Ashcraft, 1995; Ashcraft, Donley, Halas, & Vakali, 1992; Barrouillet & Fayol, 1998; Cowan, 1999; Fürst & Hitch, 2000; Kaufmann, 2002; Kaufmann, Lochy, Drexler, Semenza, 2003; Lovett, Reder, & Lebiere, 1999; Seitz & Schumann-Hengsteler, 2000, 2002). Still others are somewhere in between and suggest that retrieval is 'partially autonomous' (e.g., Zbrodoff & Logan, 1986). In the current discussion, we do not simply ask whether or not retrieval occurs automatically (i.e., whether or not it relies on cognitive resources). Instead, we ask *which stages* in direct memory retrieval might or might not require attentional resources.

When retrieving an answer from long-term memory, two processes are inevitable. First, number information in long-term memory has to be accessed; second, one answer has to be selected as the correct one. Memory access to stored number facts probably does not require any working-memory resources. In mental arithmetic, the presentation of two problem operands automatically activates the correct answer and several nodes adjacent to the correct answer (e.g., DeStefano & LeFevre, 2004; Galfano, Rusconi, & Umiltà, 2003; LeFevre et al., 1988, LeFevre & Kulak, 1994; Lemaire, Barrett, Fayol, & Abdi, 1994; Rusconi, Galfano, Speriani, & Umiltà, 2004; Rusconi, Priftis, Rusconi, & Umiltà, 2006; Thibodeau, LeFevre, & Bisanz, 1996; Zbrodoff & Logan, 1986, 1990). However, the obligatory activation of possible answers in long-term memory does not imply that these answers can be *retrieved* without any involvement of attentional resources. Instead, we believe that after the automatic activation two resource-demanding processes occur: inhibition and selection. These processes are not specific for arithmetic fact retrieval, though. Their occurrence speaks to a much broader cognitive context (e.g., retrieval of words, events, names, …). In the following, we discuss the manifestation of these processes in the context of mental arithmetic.

The presentation of two numbers of an arithmetic problem (e.g., 6 x 7) results in the automatic activation of these number nodes in long-term memory. Activation then automatically spreads from the presented nodes along associative links to related number nodes, such as the sum (13), the product (42), and table-related responses (e.g., 35, 36, 48 and 49). Hence, incorrect responses need to be *inhibited* in order not to disturb recall of relevant information (e.g., Carretti, Cornoldi, De Beni, & Palladino, 2004; Conway & Engle, 1994; Engle, Kane, & Tuholski, 1999a). Note that the inhibition point of view is in agreement with Zbrodoff and Logan (1986) arguing that retrieval is partially autonomous because it can begin without intention, but can be inhibited after its start.

Another executive function that is possibly needed in direct memory retrieval is the *selection* of relevant information (e.g., Szmalec, Vandierendonck, & Kemps, 2005). It is obvious to suppose that the process of response selection co-occurs with the inhibition process. More specifically, after the activation of possible responses in long-term memory, not only have incorrect responses to be inhibited, the correct answer has to be selected as well.

WHAT IS THE ROLE OF EXECUTIVE RESOURCES IN NONRETRIEVAL EFFICIENCY?

If one assumes that simple fact retrieval involves a cognitive cost, nonretrieval strategies should even be more demanding because they often require successive retrievals from long-term memory. The main functions of the central executive in nonretrieval strategies thus equal those needed in retrieval, i.e., response selection and inhibition. However, nonretrieval strategies incorporate still other processes that are not needed in retrieval. Each of these processes might also require cognitive resources, which explains why more executive resources are needed in nonretrieval strategies than in direct memory retrieval.

One of the main functions of the central executive is *switching* between tasks, operations, and mental sets. This switching function might be heavily needed in the transformation strategy. For example, when solving the problem 6 x 7, a participant may multiply 6 x 6 and maintain 36 in memory. Then, the participant has to switch from multiplication to addition in order to reach the correct solution: 36 + 6 = 42. Another executive function that might be needed in nonretrieval strategies is *interference control*. For example, after having solved the multiplication 6 x 6 = 36, the participant has to add 6 to 36. Working-memory resources might be needed to prevent the previous attained answer (36) from interfering with the correct answer (42). Other nonretrieval processes that might require attentional resources are keeping track of the calculation process (cf. *monitoring*) and the *integration of information* during arithmetic problem solving. Future studies are needed to test the role of these executive functions across simple-arithmetic strategies.

WHAT IS THE ROLE OF PHONOLOGICAL RESOURCES IN STRATEGY EFFICIENCY?

The role of phonological resources was investigated in Chapters 4 and 5. Retrieval was never affected by a phonological load[3]; hence, retrieval does not rely on phonological working-memory resources. This conclusion

[3] The effects of the 5-letter load on retrieval speed (cf. Chapter 4) were most certainly due to the fact that this task relied on executive resources as well.

is in agreement with brain-imaging studies in which no activation in cerebral language areas was observed during arithmetic fact retrieval (e.g., Pesenti, Thioux, Seron, & De Volder, 2000; Venkatraman, Ansari, & Chee, 2005; Zago et al., 2001) and with single-cases studies of brain-damaged patients (e.g., Whalen, McCloskey, Lindemann, & Bouton, 2002). However, we observed significant roles for both the active phonological rehearsal process and the passive phonological store when *non*retrieval strategies were used. More specifically, the active phonological rehearsal process was needed in transformation and counting strategies, whereas the passive phonological store was needed in backward counting and in the via-multiplication strategy. The possible roles of these phonological components are discussed in turn.

First, the *active phonological rehearsal process* is needed to rehearse and manipulate the number information that accesses the passive phonological store. In transformation strategies (e.g., 8 + 5 = 8 + 2 + 3 = 10 + 3 = 13), running totals have to be maintained over time in order to ensure accuracy (Logie, Gilhooly, & Wynn, 1994). In counting, the phonological rehearsal process is needed in the one-by-one incrementing process (e.g., 8 + 3 = 8… 9… 10… 11; Ashcraft, 1995). Inner speech processes, such as silent counting, also require active phonological resources (Hitch, 1978). Finally, reading out the obtained answer from the passive phonological store is a role of the active phonological rehearsal process as well (Logie & Baddeley, 1987).

A significant role for the *passive phonological store* was observed in only two cases: (1) when counting was used to solve subtraction problems (cf. Chapter 4) and (2) when the via-multiplication strategy was used to solve division problems (cf. Chapter 5). Passive phonological resources seem only to be needed when less automatic operations have to be performed (i.e., subtraction and division) and not when over-learned and almost automatic processes have to be performed (i.e., addition and multiplication). Indeed, counting-up (e.g., 7… 8… 9… 10) and reading out the multiplication tables (e.g., 7… 14… 21… 28) are over-learned processes that require very few (if any) phonological storage processes. Counting-down (e.g., 10… 9… 8… 7), on the other hand, is a process in which several numbers (e.g., the previous count and the next count) have to be maintained simultaneously, relying on the passive phonological store. The same can be said for solving divisions via multiplication: one not only has to maintain the problem operands (e.g., 42 : 6), but also the relevant multiplication problem (6 x ? = 42) while trying to infer the correct solution. Thus, the passive phonological store is needed to continually register the running total – especially when this information is not readily accessible. It should be noted, though, that the observed roles of the passive phonological store were rather limited. More effects might be found when using irrelevant speech that is phonologically similar to the numbers being counted (cf. Logie & Baddeley, 1987) or when using irrelevant speech with more spoken syllables per time unit (Buchner, Steffens, Irmen, & Wender, 1998) – ideas for future research.

In conclusion, phonological resources are needed in verbally based strategies such as transformation and counting. It is not clear, however, why a phonological load of three letters was sometimes high enough to produce interference (cf. Chapter 5), and sometimes not (cf. Chapter 4). It would be interesting to go further into this question and test from which letter load on phonological interference is created. As this load level might differ across individuals, it would be interesting to work with individually-adapted span levels.

WHAT IS THE ROLE OF WORKING MEMORY ACROSS DEVELOPMENT?

One of the main observations concerning age-related effects in strategy efficiency was that the role of working memory decreased. What causes the decreasing role of working memory across development? We see two reasons. On the one hand, due to ameliorations in strategy efficiency and strategy selection, arithmetic processing becomes more automatic and less resource demanding. These changes result in lower working-memory involvement, even if the resource pool itself does not increase with age. On the other hand, there might be a growth in the pool of cognitive resources which, in its turn, enhances arithmetic performance. Both theories do not exclude each other. Our data indicate that, alongside with an automatization in arithmetic skill, children's cognitive resources also gain in dimension. Indeed, both processing speed and working-memory span increased with age. Hence, it is the interaction between task demands and available cognitive resources that changes with age.

WHAT ABOUT THE ROLE OF WORKING MEMORY IN THEORETICAL MODELS?

The fact that executive resources are needed in efficient strategy execution is not only in agreement with the working-memory model of Baddeley and Hitch (1974), but also with many other models. We discuss three of them. According to the adaptive control of thought-rational (ACT-R) theory of Anderson (1993, 1996; Anderson et al., 1996), there is a limit on available source activation. This source activation has to be divided between all tasks at hand. When memory retrieval has to be accomplished simultaneously with another task requiring source activation, both tasks interfere with one another, resulting in poorer retrieval. Anderson and colleagues thus conclude that retrieval from long-term memory – how automatized it may be - requires the allocation of attentional resources. However, when retrieval becomes more automatic, fewer cognitive resources are needed, which explains the decreasing role of working memory across development.

The recently proposed time-based resource-sharing model of working memory (Barrouillet, Bernardin, & Camos, 2004; Barrouillet & Camos, in press; Gavens & Barrouillet, 2004; Lépine, Barrouillet, & Camos, 2005; Lépine, Bernardin, & Barrouillet, 2005) assumes that constraints on retrieval processes represent one of the cornerstones of cognitive functioning. In this model, there is a central bottleneck that precludes the simultaneous execution of two memory retrievals (see also Carrier & Pashler, 1995; Pashler, 1998; Rohrer, Pashler, & Etchegaray, 1998). In the most recent versions of the time-based resource-sharing model, the bottleneck does not only affect retrievals, but also other executive processes. Hence, each executive process that temporarily occupies the central bottleneck impedes other processes to be concurrently run. Retrieval is thus subject to interference from unrelated central processing (see also Rohrer & Pashler, 2003), resulting in less efficient retrieval use under executive loads.

Finally, the model of strategy choice and discovery simulation[4] (SCADS, Shrager & Siegler, 1998; Siegler & Arraya, 2005) includes an attentional spotlight that increases the amount of resources devoted to execution of poorly learned strategies. The more often a strategy is used, less attention will need to be focused on its execution. Although this model is primarily constructed to account for strategy selection processes, it is not in disagreement with the observation that strategy execution involves attentional resources. Moreover, it also explains the decreasing role of working memory across development.

STRATEGY SELECTION

In this section, we discuss cross-experimental findings concerning people's strategy selection (i.e., the frequency with which the different strategies are executed). Once again, the discussion is constructed upon several relevant questions that arose across the PhD project.

[4] This model is a more recent version of the adaptive strategy change model (ASCM, Siegler & Shipley, 1995).

WHY ARE PEOPLE SO RELUCTANT TO CHANGE?

As noted above, practice enhanced the frequency of retrieval use. However, participants did not easily change their strategy distributional pattern and almost never reached 100% retrieval use. The reluctance to change strategies has been shown in many cognitive domains. Already in 1942, Luchins presented results obtained in the well-known 'water jars task'. In this task, participants are presented three jar capacities (e.g., A = 21, B = 127, and C = 3) and a desired quantity (e.g., 100). They had to obtain the desired quantity by filling and pouring from the three jars. For the example above, this goal can be achieved by filling jar B from a tap, pouring from jar B to fill jar C twice, and pouring from what is left in jar B to fill jar A. Luchins observed that participants who had used this solution method for five problems in a row, 81% used the same method on test problems that could be solved far more simply (e.g., adding jar A and jar C). In contrast, control participants who received only the test problem, never used the complex solution method. Luchins concluded that participants continue to use the same solution strategy that worked well on previous problems, even when it was no longer appropriate.

Something similar might occur in solving simple-arithmetic problems. When participants find a particular strategy that works well (e.g., a reasonably fast and accurate nonretrieval strategy), they continue to use it, even when another strategy (e.g., retrieval) is more efficient. People thus often prefer a satisfactory strategy to the optimal strategy (cf. Dierckx & Vandierendonck, 2005). Indeed, learning not to use old strategies may be as challenging as learning to use new ones (cf. Kuhn, Schauble, & Garcia-Mila, 1992). However, the persistent use of nonretrieval strategies should not necessarily be seen as a maladaptive feature of human cognition; it may also be part of a larger, generally adaptive tendency toward maintaining strategic variability (cf. Siegler & Stern, 1998). Indeed, nonretrieval strategies might lose much of their utility in the context of simple arithmetic; however, they are still needed to solve complex-arithmetic problems. The reluctance to change also has implications for strategy selection models. As argued further in this section, practice should not inevitably lead to retrieval use; it might lead to the automatic activation of *non*retrieval strategies as well.

DOES WORKING MEMORY PLAY A ROLE IN STRATEGY SELECTION?

Strategy selection was not affected by working-memory loads in any study reported in the current doctoral dissertation (cf. Chapters 4, 5, and 7). Moreover, working-memory span did not predict retrieval use except for 2nd grade children (cf. Chapter 6). Thus, people's strategy selection process does not depend on working-memory resources. This result is in agreement with Hecht (2002), who observed no working-memory load effects on participants' strategy selection in an addition verification task (e.g., 5 + 8 = 12, true/false?). In a computational estimation task (e.g., 43 x 56 = ?), in contrast, Imbo, Duverne, & Lemaire (in press a) observed effects of an executive working-memory load on people's strategy selection. In the following, we offer some explanations for these different results concerning the role of working memory in the strategy selection process.

A first factor that might mediate working-memory involvement in strategy selection is *task complexity*. Obviously, estimating complex-arithmetic problems is a more complex task than verifying or solving simple-arithmetic problems. Another factor that might play a role is the *difference across the available strategies*. In simple-arithmetic problems, both retrieval and nonretrieval strategies can be used to reach the correct solution. However, direct memory retrieval is generally the dominantly used strategy. In the complex estimation task used by Imbo et al. (in press a), in contrast, each single problem had its 'best' strategy to be solved with. Hence, participants had to make a deliberate choice between the available strategies in order to choose the best strategy. The same is true in many reasoning tasks, in which one strategy can work in situation A but another strategy is needed in situation B. In such dynamic tasks, in which strategy choices have to be changed throughout the experiment, executive working-memory resources might be more heavily needed in the strategy selection process.

To conclude, we argue that working memory is not needed in choosing among several simple-arithmetic strategies. However, in other strategy domains, working-memory resources might be needed in strategy selection. In such cases, making adaptive strategy choices should consist of weighting the benefits (e.g., accuracy, speed) and the costs (e.g., load on working memory) of each strategy[5]. However, when not enough resources are available, people may fail in this cost-benefit analysis and apply one single strategy to all problems, even though this strategy is not the best one. Thus, when participants have to trade off working-memory resources between problem solving and strategy selection, both processes might be executed worse (cf. Dierckx & Vandierendonck, 2005).

WHAT SHOULD A STRATEGY SELECTION MODEL LOOK LIKE?

Multiple strategy use was observed in all our studies. Hence, there is no doubt that current selection theories need to account for multiple strategy use both in adults (Cf. Chapters 2-5) and children (cf. Chapters 6-7). Instead of developing a new model, we believe that previously developed models can readily be modified in order to account for multiple strategy use (see also Hecht, 1999, 2002; LeFevre et al., 1996a, 1996b). The ASCM (Siegler & Shipley, 1995), for example, assumes that a particular strategy is selected when that strategy can be executed efficiently. Because retrieval is generally the fastest and most accurate strategy, this strategy will be selected. Yet, when a nonretrieval strategy can be executed with reasonable speed and accuracy, this strategy also has chances to be selected. The ASCM further predicts that successful strategy execution strengthens problem-answer

---

[5] Note that both *objective* and *subjective* costs and benefits might influence one's strategy choices.

associations, which results in virtually 100% retrieval use. As many participants continue to use nonretrieval strategies, we propose a slightly modified version of the ASCM, in which problem-*strategy* association strengths are maintained relative to problem-*answer* association strengths. Nonretrieval strategies might then be automatically activated, and, consequently, initiated before retrieval. This slightly modified version of the ASCM would guarantee the continued, if infrequent, use of nonretrieval strategies.

There are, however, two possible problems with this reasoning. First, one might question whether people with strong problem-strategy associations will *ever* switch to retrieval strategies. To ensure progress in people's strategy selection process, the ASCM should incorporate two archictectural features that are included in the CMPL theory (Rickard, 1997) as well: (a) nonretrieval strategy execution strengthens retrieval nodes but not vice versa, and (b) retrievals actively inhibit nonretrieval strategies but not vice versa. Differential parameters across participants might then explain why some people are 100% retrieval users and others are not.

Second, nonretrieval strategies often need more working-memory resources than retrieval does (cf. Chapters 4 and 5). As more demanding strategies are regularly less efficient, this observation might be at variance with the assumption that nonretrieval strategies continue being used. Yet, nonretrieval strategies do *not always* need more working-memory resources. Although this is true over participants, this does not have to be true for each problem in each single participant. Suppose a person with poorly stored arithmetic facts (i.e., flat distributions of associations). This person will need many working-memory resources to retrieve arithmetic facts from memory; these working-memory resources will be involved in selecting the correct response (which might not be highly activated) and inhibiting incorrect responses (which might receive activation levels equal to the correct response). Consequently, retrieval will be very slow and resource-demanding. It is also possible that the activation of the correct answer will never reach the confidence criterion. In that case, the person might opt to choose another strategy; e.g., a nonretrieval strategy that is highly activated and automatically implemented. To conclude, in such a sperson, nonretrieval strategy execution might be *less* resource demanding than direct memory retrieval.

Note that the ASCM also offers opportunities to account for intra-individual and inter-individual differences in strategy efficiency and strategy selection. More specifically, the confidence criterion can vary from one individual to the next, and within an individual, it can vary from trial to trial. People will only state an answer if its activation level exceeds the confidence criterion. When somebody has a very rigorous confidence criterion, retrieval use decreases, even for highly activated answers. When the confidence criterion is rather lax, retrieval use will increase, even if answers are not that highly activated. Individual differences in the choice of retrieval versus nonretrieval strategies have been hypothesized to depend on the height of the individual's confidence criterion rather than on individual differences in problem-answer association strengths (Kerkman & Siegler, 1997).

WHAT ABOUT THE POSITION OF WORKING MEMORY IN STRATEGY SELECTION MODELS?

The absence of a significant role of working memory resources in strategy selection is in agreement within the ASCM (Siegler & Shipley, 1995), in which strategy selection is based on simple basic processes such as activation weighting and association strengthening and not on conscious, deliberate, or metacognitive processes requiring working-memory resources. However, as noted above, there exists evidence that – in some cases – the strategy selection process *does* need working-memory resources. This cannot be accounted for by the ASCM or by the more recent SCADS model (Shrager & Siegler, 1998; Siegler & Arraya, 2005).

However, the SCADS model includes a metacognitive system with an attentional spotlight. This attentional spotlight allocates resources to the execution of poorly learned strategies. When enough attentional resources are available, they might be used to discover new strategies or to interrupt the execution of an ongoing strategy. The attentional spotlight thus plays a role in strategy execution and strategy discovery, but not in strategy selection. Hence, future adaptations of the SCADS might think to allow the attentional spotlight to interfere with the strategy selection process in some cases (e.g., in complex-arithmetic problem solving).

ARE THERE OTHER CHALLENGES FOR FUTURE STRATEGY SELECTION MODELS?

As noted in the Introduction, the question whether retrieval and nonretrieval strategies can be *simultaneously* activated is a debated issue. The ASCM (Siegler & Shipley, 1995) and the CMPL theory[6] (Rickard, 1997) exclude parallel completion of retrieval and nonretrieval strategies. In horse race models (e.g., Logan, 1988), in contrast, retrieval and nonretrieval strategies run in parallel. There exist evidence pro and contra each type of model. Compton and Logan (1991) supported the horse race model whereas Rickard (1997) provided evidence against horse race models. He argued that because of intrinsic attentional limits only one strategy can be executed at one time. It is indeed true that, although multiple candidates for retrieval can be simultaneously activated, two retrievals cannot be completed in parallel (cf. Pashler, 1993). A recent ERP study by El Yagoubi, Lemaire, and Besson (2003) further showed that the choice between available arithmetic strategies is made within 250 msec post stimulus presentation – which also argues against parallel strategy execution. Though, we believe that it would be

[6] The CMPL model strongly resembles the ASCM. The main difference is that the CMPL model has been developed to account for adult data whereas the ASCM has been developed to account for child data.

interesting to test the horse race model against one-strategy models with ‘pure’ arithmetic tasks. As far as we know, this has not been done up until now; only alphabet arithmetic tasks (e.g., Compton & Logan, 1991; Hoyer, Cerella, & Onyper, 2003) and pseudo-arithmetic tasks (e.g., Rickard, 1997; Touron, Hoyer, & Cerella, 2004) have been used.

Another challenge for future strategy selection models is the concept of *consciousness*. Not everything in the strategy selection process needs to occur consciously. It has been shown, for example, that we have the capability to access our memories before we do a careful memory search (cf. the ‘feeling of knowing’, Reder, 1987). Consequently, strategy selection is sometimes based on the problem’s familiarity rather than on the answer’s ‘retrievability’; retrieval is then attempted as soon as the problem’s familiarity exceeds a threshold value (Metcalfe & Campbell, in press; Onyper, Hoyer, & Cerella, 2006; Reder & Ritter, 1992; Schunn et al., 1997). Hence, two types of processes might be involved in strategy selection: Strategic processes that evaluate contextual or problem-related information, on the one hand, and less conscious processes that quickly evaluate how familiar the question seems, on the other. This duality has not been implemented in strategy selection models up until now.

THE IMPACT OF INDIVIDUAL DIFFERENCES

We showed that participants’ strategy efficiency and strategy selection processes were associated with several individual differences such as gender, calculator use, and working-memory span. The possible roles of these individual differences have already been discussed in the relevant chapters (cf. Chapters 2, 5, 6, and 7). However, in the following, we again summarize the results concerning these individual differences – along with a (probably more theoretical) explanation.

ARITHMETIC SKILL

In our experiments, arithmetic skill correlated significantly with both strategy selection and strategy efficiency. Influences of arithmetic skill have been observed previously in both children and adults and can be readily explained by the fact that the mastery of basic arithmetic skills facilitates the acquisition of more complex mathematical concepts and procedures. Future research might test to what extend other individual-difference variables (e.g., daily practice) are mediated by arithmetic skill. For example, participants who are less skilled in arithmetic have been shown to choose study curricula with fewer math requirements (LeFevre, Kulak, & Heymans, 1992).

Importantly, skill-related differences across persons have been discovered in brain-imaging studies as well. In an fMRI study, Menon et al. (2000) compared adults performing perfectly on a simple-arithmetic test with adults occasionally making errors. The main difference was that perfect performers showed less activation of the left angular gyrus, indicating greater automatization and less need for rehearsal. Using the Positron Emission Tomography (PET) technique, Pesenti et al. (2001) contrasted an expert calculator with a group of non-experts. The expert did not show an increased activity in regions that also exist in non-experts; instead, he used different brain areas than did the non-experts. Finally, a recent fMRI study by So et al. (2006) showed that accurate participants had optimal arousal levels (cf. thalamus activation) and made use of working-memory resources (cf. activation in the ventrolateral prefrontal cortex). The brain activation in less accurate participants, in contrast, indicated negative emotions – possibly math anxiety (cf. activation in the middle occipital gyrus, the posterior cingulate gyrus, and the precentral gyrus).

SHORT-TERM MEMORY AND WORKING MEMORY

Effects of short-term memory span and working-memory span were tested in our developmental studies only. We observed that strategy selection and strategy efficiency were related to individual differences in working-memory span (cf. Chapter 6) but not to individual differences in short-term memory span (cf. Chapter 7), which is in agreement with other studies (e.g., Bayliss, Jarrold, Gunn, & Baddeley, 2003). Having a large working-memory span may fulfill several roles.

First, it might be needed for the simultaneous activation of the problem operands and the answer (e.g., Geary, 1993, 1994; Geary, Brown, & Samaranayake, 1991; Hecht, in press; Hecht, Torgesen, Wagner, & Rashotte, 2001). According to this theory, problem-answer associations in long-term memory can only be strengthened if both the problem and the answer can be hold simultaneously in working memory. Hence, distributions of problem-answer associations may not become sufficiently peaked in people with limited working-memory resources, resulting in less frequent and less efficient retrieval use (but see Thevenot, Barrouillet, & Fayol, 2001, for a different account). Second, individual differences in working-memory span might also reflect differences in the amount of available attentional resources (e.g., Anderson, 1993; Barrouillet & Camos, 2006; Barrouillet et al., 2004; Conway & Engle, 1994; Cowan, 1995, 1999; Engle, Tuholski, Laughlin, & Conway, 1999b; Kyllonen & Christal, 1990). Consequently, in high-span people, strategy execution does not delete the limited resource pool as much as for low-span people. According to this approach, high-span individuals will achieve higher efficiency levels even when there are no differences in the problem-answer association strengths in long-term memory. Furthermore, high-span participants would also be more able to control attention and to suppress the activation of irrelevant items (e.g., Carretti et al., 2004; Conway & Engle, 1994; Engle, 2001; Engle et al., 1999a).

Finally, note that the available working-memory capacity and the obtained arithmetic performance might influence each other. High-span participants frequently retrieve answers from long-term memory, which frees working-memory resources. Consequently, they are in a good position to devote their limited resources to other problem-solving processes. Low-span people, in contrast, often use more time-consuming and resource-demanding strategies to solve simple-arithmetic problems, which puts them at an extra disadvantage because their limited working-memory resources must then be devoted to nonretrieval strategies at the expense of other resource-demanding processes.

PROCESSING SPEED

The influence of processing speed was only tested in the developmental study reported in Chapter 7. Processing speed correlated significantly with children's strategy selection and strategy efficiency. However, when working memory was also taken into account, the predictive value of processing speed completely disappeared. This observation indicates that working memory, rather than processing speed, is an important predictor of children's arithmetic performance (see also Gavens & Barrouillet, 2004).

However, note that we did not test processing speed in any of our adult studies. Hence, future research is needed to test the role of processing speed in adults (see Duverne & Lemaire, 2004, 2005; Duverne, Lemaire, & Michel, 2003; Duverne, Lemaire, & Vandierendonck, in press, for studies in younger and older adults). Future studies might also test which subprocesses of processing speed are the most predictive. Durand, Hulme, Larkin, and Snowling (2005), for example, observed that arithmetic skill was better predicted by *digit comparison* (a subprocess of processing speed) than by general processing speed. Similarly, Hopkins and Lawson (2006) observed that *counting speed* was an important factor in explaining why practice does not always lead to retrieval. Also note that the influence of counting speed is closely related to the influence of working memory: a certain amount of counts has to be completed within the working-memory span so that the answer is achieved before the problem operands have decayed.

MATH ANXIETY

Anxiety effects were tested in Chapters 5 and 7. High-anxious participants used retrieval less frequently (cf. Chapters 5 and 7) and executed retrieval and nonretrieval strategies less efficiently (cf. Chapter 5). Note that the effects of math anxiety were not consistently significant across experiments and should thus be interpreted with caution. Moreover, the results were based on one single question only, and resulted from correlational research, which precludes making a causal connection. However, they were significant and should thus not be denied.

Ashcraft and Kirk (2001) hypothesized that math anxiety manifests itself in the form of intrusive thoughts or worries about the situation and its outcome. These thoughts and worries create a dual-task environment, in which executive resources have to be divided across the execution of the task at hand and the worrying thoughts. This working memory-based account explains why anxious persons have difficulties in strategy execution, which is resource demanding (cf. Chapters 4 and 5). Math anxiety might also have an influence during the original learning of arithmetical facts, resulting in poorly stored number facts in long-term memory. For example, math-anxious participants might want to dispose arithmetic tasks as soon as possible by sacrificing accuracy for speed, resulting in faster but more error-prone performance (cf. local avoidance, Ashcraft & Faust, 1994). The absence of correctly solved problems precludes the construction of strong problem-answer associations, resulting in slower and less frequent retrieval use.

However, math anxiety often correlates with other individual differences that might enhance the anxiety effects. For example, math-anxious persons avoid classwork and fields of study involving mathematics (cf. global avoidance, Ashcraft & Faust, 1994). The long-term avoidance of math and the lesser mastery of the math that could not be avoided might thus also affect anxious participants' arithmetic skill. Future research is needed to disentangle anxiety effects from skill effects. As participants' reports of math anxiety might be heavily biased by the willingness to uncover emotions, future studies might also think of using more objective measures of math anxiety (e.g., physiological measures).

CALCULATOR USE

Allow me to start the discussion on the influence of calculator use on people's arithmetic performance with a small story.

A question is asked to Computer Science department students.
The question is: What is the value of '2 x 2'?
The 1st year student says '4', without any thinking.
The 2nd year student says '4, exactly', after a moment of thinking.
The 3rd year student takes a pocket calculator,
presses some buttons and says '4'.
The 4th year student writes a program of about 100 lines,
debugs it, runs it and says: '4.0e+00'.
The 5th year student designs a new programming language that perfectly fits for solving such problems, implements it, writes a program, and answers:
'It says "4", but I doubt if I really fixed that ugly bug last night...'
The student just before the final graduation exams cries in desperation:
'Why, why do you think I must know all that bloody constants by heart?!'

Although this tale exaggerates the possible influences of frequent calculator use, it might contain some truth as well. The effect of calculator use was studied in Chapters 2 and 5. The frequency of calculator use influenced strategy efficiency in both chapters (i.e., more frequent calculator use induced less efficient strategy execution) and strategy selection in

Chapter 5 (i.e., more frequent calculator use reduced retrieval frequency). However, because the frequency of participants' calculator use was based on one single question only, these results should be interpreted with caution. Further research might use a more extended questionnaire to test calculator use. Another aspect that limits possible conclusions is that correlation is not causation: students poor in mental arithmetic might also be more inclined to use the calculator.

We do not think that frequent calculator use unconditionally affects people's arithmetic performance. Indeed, a recent cross-cultural study showed that, whereas students from the Netherlands use calculators quite frequently and students from Japan do not, both countries have high levels of student's arithmetic achievement (Stigler & Hiebert, 2004). We rather think that frequent calculator use is only bad when it impedes understanding. Unfortunately, the design of current pocket calculators is said to be confusing and non-mathematical (Thimbleby, 2000). Thimbleby further argues that the problem with calculators is that, unlike dictionaries, their functions are badly organized and interact confusingly with each other. Hence, it is not unthinkable that frequent use of calculators (or other devices such as cell phones) impedes conceptual understanding of arithmetic relations. It would be interesting to compare calculator use with abacus[7] use. The latter device is also an external help, but it is said to be conceptually better organized. Indeed, people using the abacus are said to have better arithmetic skills (Hatano, 2004).

[7] "Abacus" comes from the Greek word "abax" which means calculating board or calculating table. An abacus consists of a wooden frame and several rows of beads. The abacus is nowadays still used in East-Asian countries, where its use is part of the arithmetic curriculum in grade schools.

GENDER

Gender effects were investigated in Chapters 2, 5, and 7. Boys used retrieval more frequently than did girls and used retrieval also more efficiently than did girls[8]. There exist many conflicting explanations for gender effects in cognitive domains, and it is, based on the current results, impossible to know which ones are true. In the following, we list possible explanations across four broad domains: biology, personality, environment, and cognition.

A first group of explanations is *biologically* based. Examples are the hypothesis that males outperform girls in the ability to concentrate single-mindedly whereas girls outperform boys in the ability to pay attention to several topics at once (Dowker, 1996), and the hypothesis that males are object-oriented and girls people-oriented (Geary, 1996). However, there is hitherto no strong evidence that gender differences in arithmetic problem-solving are biologically based (Royer, Tronsky, Chan, Jackson, & Marchant, 1999a; Geary, 1999).

A second group of explanations is based on *personality* characteristics. In comparison with boys, girls are said to be more math-anxious and less confident (e.g., Entwisle & Baker, 1983; Felson & Trudeau, 1991), to have less positive attitudes towards math (e.g., Casey, Nuttall, & Pezaris, 1997; Catsambis, 1994; Johnson; 1984), to have lower self-concepts of mathematical ability (e.g., Andre, Whigham, Hendrickson, & Chambers, 1999; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Jacobs, 1991; Jacobs & Eccles, 1992; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Meece, Wigfield, & Eccles, 1990; Stipek & Gralinski, 1991; Vermeer, Boekaerts, & Seegers, 2000), to be less competitive (e.g., Boekaerts, Seegers, & Vermeer,

---

[8] In Chapter 7, more efficient transformation use in girls than in boys was observed. Because this effect disappeared when gender differences in processing speed were accounted for, it is not discussed further.

1995; Goldstein, 1994), and to be less impulsive and more inhibited (e.g., Davis & Carr, 2002). The relation between these personality characteristics and arithmetic performance remains to be tested explicitly, though.

Third, one might contribute gender differences to *environmental* factors. According to Beal (1994), parents and teachers expect boys to perform better in math than girls, though girls are expected to excel in language-related activities. Indeed, before having received any form of instruction, girls are generally rated higher on reading whereas boys are rated higher on mathematical ability (Lummis & Stevenson, 1990). Recently, Dar-Nimrod and Heine (2006) showed that gender differences in math performance were greatly influenced by different accounts of the origins of these differences. More specifically, women scored significantly better in an arithmetic test if they were explained that gender differences in math abilities are experiential rather than genetic. The fact that gender differences become significantly smaller after training (Verschaffel, Janssens, & Janssen, 2005) also confirms that gender differences in mathematical ability are partly due to experiential and environmental influences.

A final group of explanations is *cognitively* based. It has been suggested that boys' more elaborated spatial abilities underlie gender differences in mathematics (e.g., Geary, 1998, 1999; Geary & Burlingham-Dubree, 1989; Geary, Saults, Liu, & Hoard, 2000b). This hypothesis has received little support, however (e.g., Delgado & Prieto, 2004; Dowker, 1996; Friedman, 1995, Kimura, 1999). According to Royer et al. (1999a, 1999b), gender differences in arithmetic performance are due to a male advantage in the speed of direct fact retrieval. Recently, gender differences in general intelligence and working-memory abilities have been shown (e.g., Robert & Savoie, 2006; van der Sluis et al., 2006). More specifically, males outperform females on spatial working memory and perceptual organization while females outperform males on perceptual speed and verbal fluency.

Whether or not these differences influence arithmetic performance remains to be investigated.

To conclude, there is no consensus about the possible causes of gender differences in arithmetic performance. Anyhow, gender differences should not be exaggerated either; the range of differences in arithmetic abilities in either sex is greater than the difference between the sexes. Boys and girls are probably more alike than they are different. The main differences across individuals should probably be sought in other factors (e.g., working-memory capacity) rather than in gender.

MERITS AND LIMITS OF THE CURRENT THESIS

In the following, we discuss the merits and limits of two methodologies used in the present thesis, namely the choice/no-choice method and the retrospective strategy reportage. We end this section by discussing some alternatives that can be used instead of verbal strategy reports.

THE CHOICE/NO-CHOICE METHOD

As thoroughly argued by Siegler and Lemaire (1997), the choice/no-choice method is needed in order to obtain unbiased strategy efficiency data. In the current doctoral dissertation, we fruitfully used this method in Chapters 4, 5, and 7. However, one might ask why the choice/no-choice method was not used in *all* our experiments. There are several reasons for this.

First, we had to explore which strategies should be included in the no-choice conditions. Although people use a rich diversity of strategies to solve simple-arithmetic problems, the choice/no-choice method precludes including them all. Indeed, the choice/no-choice method is very labor-intensive and time-consuming since participants need to run at least three conditions (one choice condition and two no-choice conditions). Researchers thus have to search for a balance between practical considerations (restricting the number of strategies in the choice condition), on the one hand, and the degree of information that is needed (ecological validity), on the other (cf. Luwel, Verschaffel, Onghena, & De Corte, 2003; Luwel, Lemaire, & Verschaffel, 2005; Torbeyns, Verschaffel, & Ghesquière, 2002, 2004a, 2004b, 2005a), Hence, the Chapters in which the choice/no-choice method was *not* used (cf. Chapters 2, 3, and 6) provided us with reliable data on children's and adults' strategy choices under 'ecological' circumstances. This information could then be used in the subsequent chapters in which the choice/no-choice method *was* included.

Second, it was impossible to use the choice/no-choice method in the practice study (cf. Chapter 3). Indeed, it is unworkable to train participants while restricting their strategy use to e.g. counting. Third, the time-consuming nature of the choice/no-choice method restricted the investigation of other variables. In Chapters 2 and 6, for example, we tested the differences between addition and multiplication. Including such an operation variable in combination with the choice/no-choice method would have resulted in high-demanding experiments for both the participants and the experimenter.

Though, we acknowledge that in the current studies run *without* the choice/no-choice method, the observed effects would have been clearer *with* the choice/no-choice method. In Chapter 2, for example, we observed more efficient retrieval use in more-experienced than in less-experienced students, even though the former group used retrieval more frequently on large problems. Forcing all students to use retrieval on both small and large problems would probably have increased the group differences in retrieval efficiency. Comparably, in Chapter 6, we observed that a high working-memory span predicted higher procedural efficiency in 2^{nd} and 4^{th} graders but *lower* procedural efficiency in 6^{th} graders. As argued there, we believe that the latter result is biased by strategy selection effects (i.e., high-span 6th graders used nonretrieval strategies for the largest problems only). Running the same experiment with no-choice conditions would probably inverse this effect (i.e., high-span 6th graders would also show higher procedural efficiency).

A possible criticism on using the choice/no-choice method in combination with the selective interference paradigm is that participants in no-choice conditions have to suppress their most highly activated strategy (e.g., retrieval) in order to execute the requested strategy (e.g., transformation). This inhibition process might rely on executive working-memory resources. Hence, one might ask whether the load effects observed were due to this confound rather than to the requirements of the arithmetic strategy. To test this hypothesis, we correlated frequencies of retrieval use (under no-load) with the executive-load effect on transformation and counting latencies[9]. As expected, none of these correlations reached significance. The impact of inhibition effects should thus not be exaggerated. After all, the most activated strategy differs within as well as between individuals. To conclude, we do not believe that suppression of activated strategies is a valid explanation for the observed effects of executive working-memory load on strategy efficiencies.

Finally, note that horse race theories impede the use of the choice/no-choice method in simple-arithmetic research. Indeed, if retrieval and nonretrieval strategies run in parallel, both types of strategies take part in the race, even in no-choice conditions. Hence, as already noted above, future research is needed to test the validity of horse race models in the domain of simple arithmetic. One possibility is testing participants' performance under

---

[9] These analyses were done across participants as well as across problems. The executive-load effect was calculated as follows: transformation (or counting) RT in the executive-load condition – transformation (or counting) RT in the no-load condition.

different response deadlines. Doing so, one might infer at which point the retrieval strategy passes the nonretrieval strategy.

THE USE OF VERBAL STRATEGY REPORTS

When studying strategy selection, researchers have to know which strategies are used by the participant. In the current thesis, we decided to use self reports of strategy to obtain this information. This method often has been criticized, since such reports may be influenced by demand characteristics. Kirk and Ashcraft (2001), for example, showed that instructions biased towards nonretrieval strategy use not only changed participants' strategy selection; occasionally, it also changed their strategy efficiency. Although these results are alarming, Kirk and Ashcraft conclude that verbal strategy reports *can* provide valuable insights into the cognitive processing of mental arithmetic. Such reports must be obtained in accordance with recommended methodologies (e.g., Ericsson & Simon, 1984, 1993) and under neutral experimental instructions. Therefore, in our experiments, we provided very neutral instructions to the participants, without trying to bias them to one particular strategy. We also emphasized that the presented strategies were not meant to encourage use of a particular strategy. We neither used response deadlines since it has been shown that response time deadlines influence adult's strategy choices (Campbell & Austin, 2002).

Furthermore, in-depth analyses of our data provide evidence that the participants in the present study were veridical in their strategy reports and were probably not (or only minimally) biased by the experimenter's instructions. First, retrieval was the most frequently used strategy across all our experiments. Moreover, in each experiment, both single-strategy users as well as mixed-strategy users were present. As such, the presentation of several possible strategies did not encourage use of procedural strategies.

Reactivity to the report requirement was thus not widespread, if it occurred at all.

Second, structural variables correlated more with counting efficiency than with retrieval efficiency. In the addition experiment for example (cf. Chapter 4), the correlation between counting latencies and the minimum addend was .837 and highly significant ($p < .01$) whereas the correlation between retrieval latencies and the minimum addend was .391 and not significant ($p > .30$).

Third, in no-choice conditions (cf. Chapters 4, 5 and 7), participants had the possibility to indicate whether or not they had succeeded in using the requested strategy. Both adults and children made rather frequent use of this possibility, as the amount of non-compliant trials varied between 6% and 12%. Participants are thus capable to reflect on their past strategy execution. Further analyses showed that the amount of non-compliant trials did not differ as a function of working-memory load. Hence, removing the non-compliant trials from analyses did not influence the no-choice results.

Fourth, no-choice latencies show that participants followed the directions to use this or the other specified strategy. Across operations (cf. Chapters 4-5), mean no-choice retrieval latencies varied between 863 msec and 1040 msec, mean no-choice transformation latencies varied between 1334 msec and 2994 msec, and mean no-choice counting latencies varied between 2736 msec and 4555 msec. Similarly, choice latencies suggest that participants were able to report which strategies they had used, with mean retrieval latencies between 878 msec and 1079 msec, mean transformation latencies between 1689 msec and 2424 msec, and mean counting latencies between 1290 msec and 2890 msec. Note that the infrequent use of the counting strategy dramatically decreased counting latencies in the choice condition. Anyhow, correlations between choice latencies and no-choice latencies were extremely high for each type of strategy (each $p < .05$).

Finally, the percentages retrieval use observed in the present experiments are completely in line with results previously reported (e.g., Campbell & Timm, 2000; Campbell & Xue, 2001; Campbell et al., 2006; Geary, 1996; Geary, Bow-Thomas, Fan, & Siegler, 1993a; Geary, Frensch, & Wiley, 1993b; Geary & Wiley, 1991; Hecht, 1999, in press; LeFevre et al., 1996a, 1996b; LeFevre & Morris, 1999; Robinson, Arbuthnott, & Gibbons, 2002), in which retrieval use generally lied between 66% and 88% for addition; between 57% and 71% for subtraction; between 59% and 96% for multiplication, and between 55% and 90% for division. Note that there is a high variability across studies, which might be due to several factors such as the population studied, speed-accuracy criteria, the experimental design, and experimenter instructions. To conclude, if used correctly, self-reports can provide reliable and valid data about participants' strategy choices.

ALTERNATIVE METHODS TO INFER PARTICIPANTS' STRATEGY CHOICES

Although we acknowledge that asking participants to report which strategy they used is not without problems, we believe it is one of the best ways to infer people's strategy choices. In the following, we discuss some possible alternatives, in which no verbal strategy reports are needed. Hence, these methods avoid the potential biasing effects of self-report instructions. However, each alternative method also has one or more disadvantages.

First, one may opt to rely solely on *overt behavior* to study people's strategic behavior. It has been used by Siegler (1988b) to study children's strategy use, but he acknowledged that this method might have underestimated the frequency of (covert) nonretrieval strategies. Indeed, nonretrieval strategies can be quite fast as well (cf. Baroody, 1999) and do not always induce overt behavior. Hence, this method is not a valuable option in order to study participants' strategy use.

Second, several authors inferred strategy choices based on *response latencies* only. In alphabet arithmetic research (e.g., B + 5 = ?), latencies were supposed to depend on the magnitude of the digit addend when a counting strategy was used, whereas the magnitude of the digit addend should no longer affect performance when retrieval was used (e.g., Compton & Logan, 1991; Logan & Klapp, 1991). In 'traditional' arithmetic research, the use of multiple strategies has also been inferred from speed patterns (e.g., De Rammelaere, Stuyven & Vandierendonck, 1999, 2001; Duverne & Lemaire, 2004; Duverne et al., 2003, in press). In verification problems, latencies are larger on small-split[10] problems (e.g., 8 + 4 = 13, true/false?) than on large-split problems (8 + 4 = 21, true/false?). Consequently, for small-split problems, participants were supposed to retrieve the correct answer and then compare it with the presented answer (i.e., the exhaustive-calculation strategy), whereas for large-split problems, they were supposed to use a plausibility-checking strategy (also called the 'self-terminated verification strategy). The same reasoning has been used by testing multiplication problems that did or did not violate the five rule (e.g., 5 x 13 = 68 vs. 5 x 11 = 60; Gilles, Masse, & Lemaire, 2001; Lemaire & Reder, 1999) or that did or did not mismatch the parity rule[11] (e.g., 9 x 7 = 62 vs. 9 x 7 = 65; Lemaire & Reder, 1999).

Another alternative that is based purely on response latencies is using the *ex-Gaussian* distributional model. When applying this model to response times, one obtains *mu* and *tau* values, which refer to the mean of the normal component and the mean of the exponential component, respectively. Because the *mu* value is composed of the faster set of latencies in the distribution, it should be reflective of direct memory retrieval. The *tau* value, in contrast, is composed of the slower set of latencies in the distribution and should be reflective of procedural strategy use (cf. Campbell & Penner-Wilger, 2006; Penner-Wilger et al., 2002). The main advantage of

---

[10] The split is the difference between the presented answer and the correct answer.
[11] The five rule states that multiplication problems of which one operand equals 5 should always have a product ending with 0 or 5. The parity rule states that the product of two operands is even when at least one operand is even.

this method is that no strategy reports are needed. There is, however, at least one serious caveat in this methodology: the *mu* value does not only incorporate retrievals but also fast nonretrievals, whereas the *tau* value does not only incorporate nonretrievals but also slow retrievals. Anyhow, we believe that this method provides a good alternative to self reports, especially when the results obtained by the ex-Gaussian analyses can be compared with those obtained by analyses based on strategy reports.

Further, *priming* techniques have been used to infer which strategies people use to verify simple addition and multiplication problems. Roussel, Fayol, and Barrouillet (2002) presented the sign (+ or x) before the simple-arithmetic problem that had to be verified. It was hypothesized that procedural knowledge (cf. nonretrieval strategies) would be activated as soon as the operation sign was presented, whereas declarative knowledge (cf. retrieval) would only be activated as soon as the problem operands were presented. Because it was observed that priming the operation sign reduced verification latencies for addition but not for multiplication, Roussel et al. concluded that addition is primarily solved by nonretrieval strategies whereas multiplication would be primarily solved by direct memory retrieval. Again, the main advantage of this method is that no strategy reports are needed. A disadvantage, however, is that this method only has been used in verification tasks. The validity of this method in production tasks remains to be tested.

Thevenot and colleagues (Thevenot, Barrouillet, & Fayol, 2004; Thevenot & Oakhill, 2005, 2006) created a new paradigm, the *operand-recognition paradigm*, to infer people's strategy choices. This method has first been used in complex word problems; however, more recently, it has been used in simple mental arithmetic as well (Thevenot, Fanget, & Fayol, in press). The study of Thevenot et al. (in press) entailed two conditions. In the 'addition' condition, participants had to decide whether a third number corresponded to the sum of two previously presented numbers. In the 'comparison' condition, participants had to decide whether the third number lied between the two previously presented numbers. Afterwards, a fourth number was presented and participants had to decide whether or not this number had been presented before. Results showed that recognizing large operands was more difficult after addition than after comparison, which suggests that large problems were solved by nonretrieval strategies. Recognizing small operands, in contrast, was equally difficult after addition than after comparison, suggesting that small problems were solved by direct memory retrieval. Using the same paradigm, Thevenot, Fanget, & Fayol (2005) confirmed that retrieval is the main strategy used to solve simple addition and simple multiplication problems, whereas nonretrieval strategies would be rather frequently used to solve subtraction problems. The main advantage of this method is that no verbal strategy reports or solution latencies are needed. One of the disadvantages of this method, however, is that strategy efficiencies cannot readily be investigated.

A final alternative to the use of verbal strategy reports is the use of *eye movements*. This method has mainly been used in arithmetic word problems. Text elements that were fixated for longer were assumed to be the information on which the strategies operated (e.g., De Corte, Verschaffel, & Pauwels, 1990; Hegarty, Mayer, & Green, 1992; Hegarty, Mayer, & Monk, 1995; Verschaffel, De Corte, & Pauwels, 1992). However, eye movements have also been used during arithmetic with Arabic numbers (e.g., Suppes, 1990; Suppes, Cohen, Laddaga, Anliker, & Floyd, 1983; Verschaffel, De Corte, Gielen, & Struyf, 1994). Recently, Green, Lemaire, and Dufau (in press) examined adults' strategy choices in complex arithmetic by two means: trial-by-trial strategy reports and eye movements. The point of regard was assumed to correspond to the mental operation currently being performed. Eye movements validated participants' use of the required strategies (in no-choice conditions) and reported strategies (in choice conditions). Although this method might provide interesting possibilities for future research, one of its disadvantages is that no specific information about the strategies can be obtained. As noted by Thevenot & Oakhill (2005), with this method it is impossible to differentiate between similar but different strategies such as (65 + 2) x 4 and (4 x 65) + (4 x 2).

In conclusion, there exist several alternatives to the use of verbal strategy reports. We believe that some of them are reliable and can provide additional information about people's strategy use in mental arithmetic. However, most of them also have one or more disadvantages. Although obtaining verbal reports has disadvantages as well, we believe that it is one of the best methods to infer people's strategy choices.

PRACTICAL IMPLICATIONS

EDUCATIONAL IMPLICATIONS

We clearly showed that arithmetic abilities are based on procedural knowledge (e.g., knowing how to execute nonretrieval strategies) as well as on declarative knowledge (e.g., being able to retrieve fast and accurately from long-term memory). Also important, but less heavily stressed in the current thesis, is conceptual knowledge (e.g., understanding the principles and concepts that govern the mathematical domain). There is, however, a great debate about which type of knowledge should be most heavily stressed during schooling.

Across history, systematic drill was first seen as the best method to increase arithmetic skill (Thorndike, 1922). Knowledge of the basic arithmetic facts was purely viewed as the formation and the strengthening of individual stimulus-response associations. This method has been criticized as "mindless" and ignoring the genuine understanding of arithmetical principles, resulting in an undermining of this method for decades. Recently, Jackson and Coney (2005) argued against the undermining of rote learning. According to them, rote fact retrieval frees cognitive space and extends the number of functions that can be performed at once. The opposite view argues against the pure drill-based approach and promotes meaningful instruction and understanding (e.g., Baroody, 1984; Brownell, 1928, 1935; Lewis, 1989). The development of arithmetic skill and the acquisition of arithmetic tables was said not to proceed rote and meaninglessly, but to benefit and be facilitated by the appreciation of regularities and principles that govern them (Butterworth, Marchesini, & Girelli, 2003). Recently, Hecht, Close, & Santisi (2003) showed that conceptual knowledge uniquely contributes to the successful execution of arithmetic strategies.

We, in turn, believe that mathematical development is an iterative process whereby conceptual advances lead to strategic gains which in turn, lead to further conceptual advances (see also Byrnes, 1992; Byrnes & Wasik, 1991; Hiebert & Wearne, 1996; Rittle-Johnson & Alibali, 1999; Rittle-Johnson & Siegler, 1999; Rittle-Johnson, Siegler, & Alibali, 2001). If children need to focus all their attention on a calculation strategy, they might be less able to observe patterns in the outcomes of their problem solving (e.g., Canobi, 2005; Canobi, Reeve, & Pattison, 1998). For instance, they might not realize that the answer to the current problem (e.g., 6 + 8) is the same as the answer to the last problem they solved (e.g., 8 + 6). Children using efficient strategies are advantaged because their efficiency frees up working-memory space for considering more conceptual problem relations. This conceptual knowledge may then, in its turn, lead to the refinement of children's problem-solving skills and enhance strategy efficiency.

However, the link between conceptual knowledge and procedural competence is not always that strong: procedural competence can be achieved despite incomplete conceptual understanding (LeFevre et al., 2006) and conceptual understanding may be achieved despite severe procedural deficits (Donlan, Cowan, Newton, & Lloyd, in press). Moreover, conceptual knowledge develops slowly and is – in the beginning – situation-specific rather than abstract (Gilmore, 2006). Interestingly, it has also been shown that strategy discovery first arises at an implicit, unconscious level and only later at an explicit, conscious level (Siegler & Stern, 1998). It is thus inappropriate to think of children 'having' or 'not having' a concept or to try to determine the specific age at which children acquire different concepts or strategies.

CLINICAL IMPLICATIONS

At present, there is some general agreement about the main behavioral manifestations in mathematically-disabled children (see e.g., Geary, 2004, for a review). They have problems in the execution of nonretrieval strategies and they experience difficulties in learning, remembering, and retrieving arithmetic facts. This, in turn, contributes to the persistent use of nonretrieval strategies. In the current thesis, a large role for executive working-memory resources in normally-developing children and adults was observed (cf. Chapters 4, 5, 6, and 7). Hence, we hypothesize that deficient working-memory resources might cause mathematical problems. Previous studies provided evidence for this hypothesis (e.g., Gathercole, Alloway, Willis, Adams, 2006; Geary, 1993; Geary, Hoard, & Hamson, 1999; Hitch & McAuley, 1991; McLean & Hitch, 1999; Passolunghi, Cornoldi, & De Liberto, 1999; Passolunghi & Siegel, 2001, 2004; Passolunghi, Vercelloni, & Schadee, in press; Siegel & Ryan, 1989; Swanson, 1993; Swanson & Beebe-Frankenberger, 2004; Swanson & Sachse-Lee, 2001)[12].

Future research should test *which* working-memory functions are deficient in mathematically-disabled children. Barrouillet, Fayol, and Lathulière (1997) already showed that insufficient inhibitory mechanisms are a major cause for poor fact retrieval in mathematically-disabled

---

[12] Note that the significant role of executive working memory has not only been shown in mathematically-disabled children, but also in precocious children (e.g., Swanson, 2006).

adolescents. This was further confirmed by Bull, Johnston, & Roy (1999) and Bull and Scerif (2001), who observed that mathematically-disabled children have poorer executive resources, and especially poorer inhibition functions. Passolunghi & Pazzaglia (2005) showed that the deficit in mathematically-disabled children's working memory is not only related to inhibition processes but also to selection and memory updating processes. Finally, van der Sluis, de Jong, and van der Leij (2004) showed that mathematically-disabled children are impaired on tasks that require both inhibition and shifting rather than on pure inhibition tasks or pure shifting tasks.

The fact that mathematically-disabled children have difficulties in performing basic processes such as inhibiting and switching might thus explain why these children have difficulties in the execution of both retrieval and nonretrieval strategies. Future research might use the approaches adopted in the current doctoral dissertation (e.g., the choice/no-choice method and the selective interference paradigm) to investigate the specific problems of mathematically-disabled children more deeply.

RESEARCH IMPLICATIONS

Finally, we would like to point to some implications for experimental settings. It is generally known that most experiments on mental arithmetic are carried out on university psychology students. When only this group is studied, results can be biased, since the student flow from secondary school to university is not equally divided across the students' amounts of arithmetic experience. Less-experienced students often choose to study psychology whereas more-experienced students often choose to study exact sciences. Percentages of male and female students differ across study branches as well, with extremely high amounts of female students studying psychology. Further, when running experiments said to test 'arithmetic skill', math-anxious people might not want to participate (cf. self-selection).

These factors imply that experiments testing solely psychology students might be biased. Therefore, researchers should be careful when generalizing their results to the total population. Future studies might think (1) to test university students of different branches, (2) to include equal amounts of males and females, and (3) to be careful when describing the topic-to-be-tested.

AVENUES FOR FURTHER STUDIES

EMOTION AND MOTIVATION

Although the investigation of individual differences is interesting, it poses several problems. One of the main problems is that *inter*-individual differences cannot be experimentally manipulated. Consequently, results obtained in individual-difference studies risk to be biased by confounding variables such as general intelligence, past experiences, gender, et cetera. However, future research might try to manipulate *intra*-individual variables, such as emotion and motivation, in order to gain more insight in the role of these variables.

For example, future research might opt to manipulate people's anxiety level. This has first been tried by Hopko et al. (2003), who utilized 7% carbon dioxide (CO_2) gas to induce anxiety. Physiological data confirmed that participants in the CO_2 group experienced more autonomic arousal; however, the gas did not affect their arithmetic performance. Math anxiety (as measured by a questionnaire), in contrast, did affect their arithmetic performance. More recently, Beilock, Kulp, Holt and Carr (2004) studied arithmetic performance under pressure. Pressure was created by monetary incentives, peer pressure, and social evaluation. Results showed that this manipulation was successful: participants in the high-pressure group showed higher levels of anxiety, felt more pressure to perform at a high level, and had worse perceptions of their performance than did participants in the low-pressure group. Importantly, pressure harmed performance only on difficult, unpracticed problems which relied heavily on working memory. When problems were practiced until their answers were directly retrieved from long-term memory, no effects of pressure were observed. It may be interesting to use a comparable manipulation as used by Beilock and colleagues, in combination with the choice/no-choice method, in order to infer whether anxiety affects strategy efficiency, strategy selection, or both.

Math-related feelings do not always have to be negative, though. LeFevre et al. (1996a) reported a significant relationship between positive attitudes toward math and direct retrieval use. More recently, Lepola, Niemi, Kuikka and Hannula (in press) observed that, from preschool onwards, motivational orientations made unique contributions to subsequent arithmetic performance. Hence, future studies may focus on the effects of such reinforcements as well.

COMPLEX-ARITHMETIC STRATEGIES

Very few studies investigated strategies in complex-arithmetic performance. The great variety in possible strategies may be responsible for this. The problem 20 x 39, for example, can be solved as follows: (20 x 40) – 20 = 800 – 20 = 780 or as follows: (20 x 30) + (20 x 9) = 600 + 180 = 780. Moreover, complex-arithmetic strategies might differ across individuals (e.g., Torbeyns, Verschaffel, & Ghesquière, 2005b) and across cultures (e.g., Shanahan, Lucidi, LeFevre, & Cestari, 2005).

Note that complex-arithmetic problems do not always have to be solved exactly. In many daily situations, estimation strategies might provide a sufficiently accurate answer. One may, for example, estimate that 21 x 39 should more or less equal 800 (cf. 20 x 40). Several studies investigated such estimation strategies (e.g., Dowker, 1997; Lemaire, Arnaud, & Lecacheur, 2004; Lemaire & Lecacheur, 2002; Lemaire, Lecacheur, & Farioli, 2000; Lemaire & Machard, 2003; Levine, 1982). Recently, Imbo et al. (in press a) showed that both the efficiency and the selection of estimation strategies are affected by an executive working-memory load. As it has been shown that exact complex-arithmetic solving is even more resource-demanding than approximate complex-arithmetic solving (Kalaman & LeFevre, in press), it might be interesting to test whether the results obtained by Imbo and colleagues still hold (or are even boosted) when complex-arithmetic problems have to be solved exactly rather than approximately.

Another complex-arithmetic process that requires working-memory resources is carrying[13] (e.g., Imbo, Vandierendonck, & De Rammelaere, in press e; Imbo, Vandierendonck, & Vergauwe, in press g; Fürst & Hitch, 2000; Noël, Désert, Aubrun, & Seron, 2001). As there are several strategies to solve carry problems, it might be interesting to test whether these strategies differently involve working-memory resources. In complex-arithmetic problem solving, Hitch (1978) observed that starting by adding the units was used more frequently than starting by adding the hundreds. However, some people changed their strategy choices when carrying was needed. These people started by adding the units when carrying was required, but started by adding the hundreds when no carrying was required. By using the choice/no-choice method, Green et al. (in press) confirmed that starting by adding the hundreds was more efficient on no-carry problems whereas starting by adding the units was more efficient on carry problems. Future research is needed to test the role of the different working-memory components across these strategies. An additional variable that might be incorporated in such studies is the presentation format. Indeed, Trbovich and LeFevre (2003) hypothesized that starting by adding the units would be used in vertically presented problems rather than in horizontally presented problems. However, because Trbovich and LeFevre obtained no strategy reports, this hypothesis still needs to be confirmed empirically.

[13] Carrying is needed when the sum of the units/tens/hundreds/… crosses 10. In the problem 526 + 138, for example, the sum of the units exceeds 10, which means that a “1” as to be carried to the tens.

THE ROLE OF OTHER WORKING-MEMORY COMPONENTS

In the current thesis, we investigated the role of phonological and executive working-memory resources in adults (cf. Chapters 4 and 5), and the role of executive working-memory resources in children (cf. Chapters 6 and 7). Further research might use the combination of the selective interference paradigm and the choice/no-choice method to investigate (1) the role of visuo-spatial working-memory resources in adults' strategy use, and (2) the role of visuo-spatial and phonological working-memory resources in children's strategy use.

Concerning the first issue, it has been hypothesized that Chinese-speaking participants store and access number facts by using phonological codes because of educational factors and the structure of their number language (LeFevre, Lei, Smith-Chant, & Mullins, 2001). European participants may rather use visual or abstract number codes. However, a recent fMRI study by Burbaud et al. (2000) showed that the involvement of phonological and visuo-spatial working-memory resources differs across individuals – rather than across cultures (see also Sohn et al., 2004). More specifically, in participants relying on a verbal strategy the main brain activation was located in the left dorsolateral frontal cortex. In participants using a visual strategy, in contrast, a bilateral activation in the prefrontal cortex and a high activation in the left inferior parietal cortex were observed. Future research might test the inter-individual and inter-cultural differences in the involvement of phonological and visuo-spatial working-memory resources.

Concerning the second issue, we hypothesize that the role of the different working-memory components would vary as a function of age. Based on our results that nonretrieval strategies require phonological working-memory resources (cf. Chapters 4 and 5) and that nonretrieval strategy use decreases with age (cf. Chapters 6 and 7), we hypothesize that the role of phonological working-memory resources will decrease with age.

In a longitudinal study, Hecht et al. (2001) confirmed that phonological working-memory abilities play a great role in 2nd to 5th graders' arithmetic skill. Moreover, the influence of some of these phonological processes was limited to 2nd and 3rd graders only. More recently, it has been shown that even 1st graders' arithmetic performance is significantly predicted by phonological resources (e.g., Fuchs et al., 2005, 2006; Noël, Seron, & Trovarelli, 2004). Finally, Grube (2005) showed that a phonological working-memory load affected arithmetic performance only for young children and not for older children. All these studies suggest that the role of phonological working memory decreases with age. However, this hypothesis remains to be tested empirically (e.g., by using strategy reports and phonological working-memory loads).

Because children's counting-based strategies often rely on concrete representations, we hypothesize that the role of visuo-spatial working memory in simple-arithmetic strategies would decrease with age as well. De Smedt, Ghesquière, and Verschaffel (2004) showed that visuo-spatial working memory was an important predictor of 1st graders' arithmetic performance, whereas phonological working memory was an important predictor of 5th grader's arithmetic performance. Central executive resources, in contrast, predicted both 1st and 5th graders' arithmetic performance. Similarly, Rasmussen and Bisanz (2005) showed that the best and only predictor of preschool children's arithmetic performance was their visuo-spatial working memory. Finally, Kroesbergen and Van Luit (2005) confirmed that individual differences in young children's arithmetic performance are predicted by visuo-spatial working memory rather than by phonological or executive working memory. Future research is needed to specify in which strategies these age-related changes in visuo-spatial involvement occur.

THE ROLE OF THE DIFFERENT EXECUTIVE FUNCTIONS

In the current doctoral dissertation (cf. Chapters 4, 5, and 7), we showed that executive resources are needed in simple-arithmetic problem solving. Using nonretrieval strategies even involved more executive resources than retrieval did. However, because the executive secondary task used in the current thesis taxed several executive components (e.g., attentional control, coordination of information, inhibition of irrelevant information, response decision), it was impossible to know *which* specific executive functions were needed. Hence, future research is needed to test whether various executive working-memory functions are differentially needed across the strategies. Such research needs to compare arithmetic performance under various load conditions (Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Coeman, 2007; Deschuyteneer, Vandierendonck, & Muyllaert, 2006).

We hypothesize that *inhibition* will play a great role in both retrieval and nonretrieval strategies; e.g., to inhibit highly activated but incorrect responses and to inhibit no-longer relevant information after having performed an intermediate step, respectively. In children, inhibition has already been shown to predict arithmetic competency (Espy et al., 2004). Another executive function that might be investigated is *switching* or *shifting*. Indeed, adaptive task performance often requires frequent (trial-by-trial) switches across strategies. Deficient strategy shifting might heavily impair performance. In a noun-pair learning task, it has been shown that performance improvements were determined by efficient strategy shifting rather than by efficient retrieval or nonretrieval strategies (Touron & Hertzog, 2004). The issue of strategy shifting in mental arithmetic still needs to be discovered, though (but see Luwel, Bulté, Onghena, & Verschaffel, 2007).

STRATEGY ADAPTIVITY

Strategies are adaptively chosen if they are the best (i.e., fastest and most accurate) strategy to solve the presented problem. An expert is thus not someone who always uses the 'one best strategy'; an expert is rather someone who uses many strategies and adapts his/her strategy choices to the problems and the situation. Applied to mental arithmetic, this reasoning indicates that frequent memory retrieval does not, in itself, indicate high mathematical ability or adaptive strategy use; it might also indicate a rather lenient confidence criterion (cf. Siegler, 1988a).

Strategy adaptivity was not tested in the current thesis. Hence, future research is needed to investigate strategy adaptivity in the domain of mental arithmetic. More specifically, it would be interesting to test the role of working memory and the influence of individual differences in strategy adaptivity. Recent studies already started to address some of these points. Gilles et al. (2001), for example, showed that strategy adaptivity was higher in high-skill than in low-skill participants. Imbo et al. (in press a) observed lower levels of strategy adaptivity under an executive working-memory load. Another interesting question concerns the development of strategy adaptivity. Recent studies already indicated that strategy adaptivity first (i.e., between childhood and adulthood) increases (e.g., Luwel et al., 2005) and then (i.e., between young and old adulthood) decreases again (e.g., Duverne & Lemaire, 2004, 2005; Duverne et al., 2003; Green et al., in press; Lemaire et al., 2004). Finally, strategy adaptivity might also be used to study mathematical disabilities. Torbeyns et al. (2002, 2004a, 2004b) already showed that normally-developing children make more adaptive strategy choices than do mathematically-disabled children.

To conclude, there is space for more research in the domain of arithmetic strategies. As becoming for scientific research, the current doctoral dissertation did not only answer relevant questions, it also raised new questions…

NEDERLANDSTALIGE SAMENVATTING

INLEIDING

In het dagelijkse leven worden we vaak geconfronteerd met getallen en cijfers waarop we één of andere bewerking moeten uitvoeren (e.g., 6 x 7, 8 + 5). Bij dergelijke eenvoudige rekenopgaven kunnen de meeste volwassenen de juiste oplossing meteen 'ophalen' uit hun lange-termijn geheugen. Dat wil zeggen, ze weten meteen dat 6 x 7 gelijk is aan 42 en dat 8 + 5 gelijk is aan 13. Dit is echter niet *altijd* zo. Ongeveer tien jaar geleden werd onomstotelijk aangetoond dat zelfs volwassenen niet alle oplossingen onmiddellijk 'uit het hoofd' weten (LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). Men gebruikt dus ook andere (procedurele) strategieën, zoals transformeren (i.e., een strategie waarbij één of meerdere tussenstappen gemaakt worden, e.g., 6 x 7 = 6 x 6 + 6 = 36 + 6 = 42) en tellen (e.g., 8 + 5 = 8… 9 … 10… 11… 12… 13). Dit zijn strategieën die vaak ook door kinderen gebruikt worden als ze leren hoofdrekenen. Deze strategieën worden doorheen de ontwikkeling echter grotendeels (maar dus niet helemaal) vervangen door ophaling uit het lange-termijn geheugen.

Omdat lange tijd gedacht werd dat volwassenen enkel de 'ophalings-strategie' gebruikten, heeft het onderzoeksdomein naar strategieën in hoofdrekenen zich pas de laatste jaren echt ontwikkeld (voor een overzicht, zie Hoofdstuk 1). Niettegenstaande er dus redelijk wat geweten was omtrent hoofdrekenen 'in het algemeen', was de kennis omtrent reken*strategieën* eerder schaars. Met deze thesis werd geprobeerd deze leemte enigszins op te vullen. In zes empirische hoofdstukken wordt de relevantie van strategieën in rekenonderzoek verder uitgespit. Er worden vier brede thema's behandeld: (1) De rol van oefening en ervaring, (2) De rol van het werkgeheugen, (3) De ontwikkeling van rekenstrategieën, en (4) Het belang van individuele verschillen.

METHODE

Er werden in elke studie twee belangrijke vragen gesteld. De eerste vraag heeft betrekking op de *strategieselectie*: Welke strategieën worden gekozen om rekenopgaven op te lossen? Om dit te achterhalen werd gebruik gemaakt van verbale strategierapportage. Dit wil zeggen dat de participanten zelf moesten aangeven hoe ze de rekenopgave hadden opgelost. Ze kregen hiervoor de keuze uit 'Onthouden' (i.e., het antwoord ophalen uit het lange-termijn geheugen), 'Transformeren', 'Tellen', of 'Iets anders' (e.g., gokken of meerdere strategieën combineren). Alhoewel er kritiek is op deze verbale rapporteringsmethode (e.g., Kirk & Ashcraft, 2001), geeft ze – mits verantwoord gebruik – wel degelijk betrouwbare en valide informatie over de strategieselectie.

De tweede vraag heeft betrekking op de *strategie-efficiëntie*: Hoe snel en hoe accuraat wordt de gekozen strategie uitgevoerd? Om de snelheid te meten, werd gebruik gemaakt van een voice key. Dit is een kleine microfoon die (tot op 1 milliseconde nauwkeurig) getriggerd wordt van zodra de participant iets zegt. De accuraatheid werd online door de proefleider bijgehouden.

Om valide data te verkrijgen over zowel strategieselectie als strategie-efficiëntie werd in sommige hoofdstukken gebruik gemaakt van de keuze/geen-keuze methodiek (Siegler & Lemaire, 1997). Volgens deze methodiek worden de participanten in twee verschillende condities getest. In de *keuze* conditie mogen ze kiezen welke strategie ze gebruiken om de rekenopgaven op te lossen. De keuzes die hier gerapporteerd worden geven een indicatie van de strategieselectie. Daarnaast nemen alle participanten ook deel aan verschillende *geen-keuze* condities. In deze condities wordt er van hen verwacht dat ze één en dezelfde strategie gebruiken om alle aangeboden rekenopgaven op te lossen. De snelheden en accuraatheden die in deze condities opgemeten worden, geven valide strategie-efficiëntie data.

RESULTATEN

De rol van oefening. Vorig onderzoek toonde reeds aan dat ervaren rekenaars sneller rekenden dan minder ervaren rekenaars. Hoe dit precies kwam, was echter onduidelijk: Had dit te maken met een betere strategieselectie (e.g., een frequenter gebruik van de ophalings-strategie) of met een hogere strategie-efficiëntie (e.g., sneller kunnen uitvoeren van bepaalde strategieën)? Om deze vraag te onderzoeken, werden verschillende methodes toegepast. In Hoofdstuk 2 werd een 'ecologische' methode gebruikt. In plaats van zelf de mate van ervaring te gaan manipuleren, onderzochten we middelbare-school studenten uit verschillende studierichtingen (i.e., met veel en weinig wiskunde lessen). We observeerden dat studenten met meer wiskunde-ervaring de oplossing op rekenopgaven vaker 'uit het hoofd' wisten. Bovendien waren ze ook sneller in het uitvoeren van zowel ophalings- als procedurele strategieën.

In Hoofdstuk 3 werd de hoeveelheid oefening gemanipuleerd: 1e bachelor studenten werden getraind in het oplossen van rekenopgaven en vervolgens getest. Ook hier werd vastgesteld dat de mate van oefening een positief effect had op zowel strategieselectie als strategie-efficiëntie. Opvallend is echter dat – in beide studies – de effecten sterker waren voor de vermenigvuldiging dan voor de optelling. Bovendien waren de effecten niet altijd zo sterk als verwacht. Men verandert blijkbaar niet graag van strategie als men reeds een strategie heeft die goed werkt.

De rol van het werkgeheugen. In het werkgeheugen wordt informatie voor een korte termijn bijgehouden en verwerkt. Volgens Baddeley en Hitch (1974) kan het werkgeheugen verder opgedeeld worden in drie delen: de centrale verwerker, de fonologische lus, en het visuo-spatiaal schetsblad. De fonologische lus kan verder opgesplitst worden in een passieve opslagplaats en een actief herhalingsproces. Vorig onderzoek toonde reeds aan dat de centrale verwerker een grote rol speelt in hoofdrekenen. De fonologische lus daarentegen, zou slechts in een beperkt aantal omstandigheden nodig zijn (zie DeStefano & LeFevre, voor een review). Om de rol van het werkgeheugen te bestuderen werd gebruik gemaakt van het selectieve interferentie paradigma, waarbij elke participant getest wordt in een enkele-taak conditie (zonder werkgeheugenbelasting) en een dubbeltaak conditie (waarin selectief één werkgeheugencomponent belast wordt).

De rol van het werkgeheugen werd onderzocht in optellings- en aftrekkingsstrategieën in Hoofdstuk 4, en in vermenigvuldigings- en delingsstrategieën in Hoofdstuk 5. In beide studies namen 1^{e} bachelor studenten deel. De resultaten toonden aan dat de centrale verwerker nodig was in de uitvoering van zowel ophalings- als procedurele strategieën. Het actief fonologische herhalingsproces was enkel nodig in procedurele strategieën. De passieve opslagplaats, tot slot, was enkel nodig wanneer procedurele strategieën gebruikt werden in de minder geautomatiseerde operaties zoals aftrekken en delen. Geen enkele werkgeheugencomponent speelde een significante rol in strategieselectie.

De ontwikkeling van rekenstrategieën. Waar procedurele strategieën slechts af en toe gebruikt worden door volwassenen, gebruiken kinderen deze relatief frequent. Kinderen leren namelijk rekenen aan de hand van telstrategieën (e.g., 4 + 3 = 4… 5… 6… 7; 3 x 8 = 8… 16… 24). Naarmate ze ouder worden en meer scholing genieten, schakelen ze over naar de ophalings-strategie. In Hoofdstuk 6 werd onderzocht in welke mate het strategiegebruik van kinderen overeenstemt met dat van volwassenen. In deze studie namen kinderen uit het 2^{e}, 4^{e}, en 6^{e} leerjaar deel. Er werd aangetoond dat het probleemgrootte-effect (i.e., grote opgaven zoals 7 x 8 worden efficiënter opgelost dan kleine opgaven zoals 3 x 4) kleiner werd naarmate kinderen ouder worden; dit was het gevolg van zowel veranderingen in de strategieselectie als in de strategie-efficiëntie. Daarnaast werd ook geobserveerd dat deze veranderingen eerst (i.e., tussen het 2^{e} en het 4^{e} leerjaar) plaatsgrijpen in de vermenigvuldiging, en pas later (i.e., tussen het 4^{e} en het 6^{e} leerjaar) in de optelling. Dit heeft waarschijnlijk te maken met de grote nadruk die gelegd wordt op het memoriseren van de tafels van vermenigvuldiging. Tot slot werd geobserveerd dat kinderen met een grote werkgeheugenspan bevoordeeld zijn ten opzichte van kinderen met een kleine werkgeheugenspan; dit gold voor strategieselectie als voor strategie-efficiëntie. De voordelen die gepaard gingen met het hebben van een grote werkgeheugenspan daalden wel met de leeftijd.

De rol van het werkgeheugen tijdens de ontwikkeling werd verder onderzocht in Hoofdstuk 7. Deze keer werd echter gebruik gemaakt van het selectieve interferentie paradigma, dat reeds gebruikt werd in de volwassenenstudies (cf. Hoofdstukken 4 en 5). De participanten in deze studie waren kinderen van het 4^e, 5^e, en 6^e leerjaar; de onderzochte operatie was de optelling. De ophalings-strategie en de telstrategie werden efficiënter uitgevoerd naarmate de kinderen ouder werden. Voor deze strategieën werd de impact van de secundaire taak (een werkgeheugenlading) ook kleiner met de leeftijd. De transformatiestrategie daarentegen, werd even efficiënt uitgevoerd door de oudere als door de jongere kinderen – de impact van de werkgeheugenlading daalde dan ook niet met de leeftijd. Net zoals bij volwassenen speelde het werkgeheugen geen rol in de strategieselectie.

Het belang van individuele verschillen. In Hoofdstukken 2, 5, 6 en 7 werd de rol van enkele individuele kenmerken onderzocht. Een variabele die altijd sterk verbonden was met zowel strategieselectie als strategie-efficiëntie was algemene *rekenvaardigheid*. Dit is niet verwonderlijk, aangezien het efficiënt kunnen oplossen van eenvoudige rekenopgaven zoals 3 x 9 een vereiste is voor het oplossen van complexere rekenopgaven zoals 3 x 89. Verder werd geobserveerd dat een frequent *rekenmachinegebruik* negatief correleerde met de participanten hun rekenprestaties. Ook *wiskundeangst* bleek niet bevorderlijk. Het hebben van een grote *verwerkingssnelheid* en/of een grote *werkgeheugenspan* was dan weer positief gerelateerd aan efficiënt strategiegebruik. In sommige studies werden *geslachts*verschillen geobserveerd. Mannelijke participanten gebruikten de ophalings-strategie vaker dan vrouwelijke participanten;

mannen hadden soms ook een efficiëntere strategie-uitvoering dan vrouwen. Aangezien het onderzoek naar individuele verschillen gebaseerd was op correlationele analyses, is het echter ongeoorloofd uit deze resultaten causale verbanden af te leiden.

DISCUSSIE

In Hoofdstuk 8 worden de belangrijkste resultaten eerst nog eens kort samengevat. Daarna wordt er dieper ingegaan op enkele relevante vragen – zowel met betrekking tot strategie-efficiëntie als strategieselectie. Er wordt stilgestaan bij de pluspunten en minpunten van deze thesis, alsook bij de praktische implicaties. Tot slot worden een aantal ideeën voor verder onderzoek gegeven.

REFERENCES

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Ackerman, P.L., & Cianciolo, A.T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences in skill acquisition. *Journal of Experimental Psychology: Applied, 6*, 259-290.

Adams, J.W., & Hitch, G.J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*, 21-38.

Alexander, L., & Martray, C. (1989). The development of an abbreviated version of the mathematics anxiety rating-scale. *Measurement and Evaluation in Counseling and Development, 22, 143-150.*

Anderson, J.R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J.R. (1996). *A simple theory of complex cognition*. American Psychologist, 51, 355-365.

Anderson, J.R., Reder, L.M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology, 30*, 221-256.

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching, 36*, 719-747.

Ansari, D., Garcia, N., Lucas, E., Hamon, K., Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport, 16*, 1769-1773.

Ashcraft, M.H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2*, 213-236.

Ashcraft, M.H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C.J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology: Progress in cognitive development research* (pp. 302-338). New York: Springer-Verlag.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106.

Ashcraft, M.H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition, 1*, 3-34.

Ashcraft, M.H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*, 181-185.

Ashcraft, M.H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 527-538.

Ashcraft, M.H., Donley, R.D., Halas, M.A., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 301-329). Amsterdam: Elsevier.

Ashcraft, M.H., & Faust, M.W. (1994). Mathematics anxiety and mental arithmetic performance: An exploratory investigation. *Cognition & Emotion, 8*, 97-125.

Ashcraft, M.H., & Fierman, B.A. (1982). Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology, 33*, 216-234.

Ashcraft, M.H., & Kirk, E.P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224-237.

Ashcraft, M.H., & Stazyk, E.H. (1981). Mental addition: A test of 3 verification models. *Memory & Cognition, 9*, 185-196.

Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A.D. (1990). *Human memory*: Theory and practice. Oxford: Oxford University Press.

Baddeley, A.D. (1992). Working memory. *Science, 256*, 556-559.

Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.

Baddeley, A.D., & Hitch, G.J. (1974). Working Memory. In G. Bower (Ed.). *The Psychology of Learning and Motivation* (pp. 47-90). New York: Academic Press.

Baddeley, A.D., & Logie, R.H. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery* (pp. 179-197). Hillsdale, NJ: Erlbaum.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multi-component model. In A. Miyake & P. Shah (Ed.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York: Cambridge University Press.

Baddeley, A.D., & Salamé, P. (1986). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 525-529.

Baroody, A.J. (1983). The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic. *Developmental Review, 3*, 225-230.

Baroody, A.J. (1984). A re-examination of mental arithmetic models and data: A reply to Ashcraft. *Developmental Review, 4*, 148-156.

Baroody, A.J. (1985). Mastery of the basic number combinations: Internalization of relationships or facts? *Journal of Research in Mathematics Education, 16*, 83-98.

Baroody, A.J. (1994). An evaluation of evidence supporting fact-retrieval models. *Learning and Individual Differences, 6*, 1-36.

Baroody, A.J. (1999). The roles of estimation and the commutativity principle in the development of third graders' mental multiplication. *Journal of Experimental Child Psychology, 74*, 157-193.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*, 83-100.

Barrouillet, P., & Camos, V. (2006). *Attention and long-term memory in the time-based resource-sharing model.* Paper presented at the 3rd European Working Memory Symposium, Genova, Italy.

Barrouillet, P., & Camos, V. (in press). The time-based resource-sharing model of working memory. In N. Osaka R. Logie (Eds.), *Working memory: Behavioral and neural correlates*. Oxford University Press.

Barrouillet, P., & Fayol, M. (1998). From algorithmic computing to direct retrieval: Evidence from number and alphabetic arithmetic in children and adults. Memory & *Cognition, 26*, 355-368.

Barrouillet, P., Fayol, M., & Lathulière, E. (1997). Selecting between competitors in multiplication tasks : An explanation of the errors produced by adolescents with learning difficulties. *International Journal of Behavioral Development, 21*, 253-275.

Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*, 183-204.

Bayliss, D.M., Jarrold, C., Gunn, D.M., & Baddeley, A.D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, 132, 71-92.

Beal, C.R. (1994). *Boys and girls: The development of gender roles*. NY: McGraw-Hill.

Beilock, S.L., Kulp, C.A., Holt, L.E., Carr, T.H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General, 133*, 584-600.

Bisanz, J., Morrison, F.J., and Dunn, M. (1995). Effects of age and schooling on the acquisition of elementary quantitative skills. *Developmental Psychology, 31*, 221-236.

Boekaerts, M., Seegers, G., & Vermeer, H. (1995). Solving math problems: Where and why does the solution go astray? *Educational Studies in Mathematics, 28*, 241-262.

Brigman, S., & Cherry, K.E. (2002). Age and skilled performance: Contributions of working memory and processing speed. *Brain and Cognition, 50*, 242-256.

Brownell, W.A. (1928). *The development of children's number ideas in the primary grades*. Chicago, IL: The University of Chicago Press.

Brownell, W.A. (1935). Psychological considerations in the learning and teaching of arithmetic. In W.D. Reeve (Ed.), *The teaching of arithmetic* (10^{th} Yearbook of the National Council of Teachers of Mathematics, pp. 1-31). New York: Columbia University, Teachers College, Bureau of Publications.

Buchner, A., Steffens, M., Irmen, L., & Wender, K.F. (1998). Irrelevant auditory material affects counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 48-67.*

Bull, R., & Johnston, R.S. (1997). Children's arithmetical difficulties: contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology, 65*, 1-24.

Bull, R., Johnston, R.S., & Roy, J.A. (1999). Exploring the roles of the visuo-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*, 421-442.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273-293.

Burbaud, P., Camus, O., Guehl, D., Bioulac, B., Caillé, J.-M., & Allard, M. (2000). Influence of cognitive strategies on the pattern of cortical activation during mental subtraction. A functional imaging study in human subjects. *Neuroscience Letters, 287*, 76-80.

Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.

Butterworth, B., Marchesini, N., & Girelli, L. (2003). Basic multiplication combinations: Passive storage or dynamic reorganization? In A. Baroody & A. Dowker (Eds.), *The development of arithmetical concepts and skills* (pp. 189-202). Mahwah, NJ: LEA.

Byrnes, D. (1992). Addressing Race, Ethnicity, and Culture in the Classroom. In D. Byrnes and G. Kiger (Eds.), *Common bonds: Antibias teaching in a diverse society*. Wheaton, MD: Association for Childhood Education International.

Byrnes, D., & Wasik, B.A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology, 27*, 777-786.

Camos, V., & Barrouillet, P. (2004). Adult counting is resource demanding. *British Journal of Psychology, 95*, 19-30.

Campbell, J.I.D. (1987a). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*, 109-123.

Campbell, J.I.D. (1987b). Production, verification, and priming of multiplication facts. *Memory & Cognition, 15*, 349-364.

Campbell, J.I.D. (1994). Architectures for numerical cognition. *Cognition, 53*, 1-44.

Campbell, J.I.D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121-164.

Campbell, J.I.D. (1997). On the relation between skilled performance of simple division and multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 1140-1159.

Campbell, J.I.D. (1999). Division by multiplication. *Memory & Cognition, 27*, 791-802.

Campbell, J.I.D., & Austin, S. (2002). Effects of response time deadlines on adults' strategy choices for simple addition. *Memory & Cognition, 30*, 988-994.

Campbell, J.I.D., & Clark, J.M. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 920-929.

Campbell, J.I.D., & Clark, J.M. (1992). Cognitive number processing: An encoding-complex perspective. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 457-491). Amsterdam: Elsevier Science Publishers.

Campbell, J.I.D., & Fuchs-Lacelle, S., & Phenix, T.L. (2006). Identical elements model of arithmetic memory: Extension to addition and subtraction. *Memory & Cognition, 34*, 633-647.

Campbell, J.I.D., & Fugelsang, J. (2001). Strategy choice for arithmetic verification: Effects of numerical surface form. *Cognition, 80*, B21-B30.

Campbell, J.I.D., & Graham, D.J. (1985). Mental multiplication skill – Structure, process, and acquisition. *Canadian Journal of Psychology, 39*, 338-366.

Campbell, J.I.D, & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition, 86*, 71-96.

Campbell, J.I.D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In J.I.D. Campbell (Ed.), *The nature and origin of mathematical skills*, (pp. 331-364). Amsterdam: Elsevier Science Publishers.

Campbell, J.I.D., Parker, H.R., & Doetzel, N.L. (2004). Interactive effects of numerical surface form and operand parity in cognitive arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 51-64.

Campbell, J.I.D., & Penner-Wilger, M. (2006). Calculation latency: The mu of memory and the tau of transformation. *Memory & Cognition, 34*, 217-226.

Campbell, J.I.D., & Tarling, D.P.M. (1996). Retrieval processes in arithmetic production and verification. *Memory & Cognition, 24*, 156-172.

Campbell, J.I.D., & Timm, J. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin & Review, 7*, 692-699.

Campbell, J.I.D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*, 299-315.

Canobi, K.H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology, 92*, 220-246.

Canobi, K.H., Reeve, R.A., Pattison, P.E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34*, 882-891.

Carpenter, T.P., & Moser, J.M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education, 19*, 345-357.

Carr, M. (1996). Metacognitive, motivational, and social influences on mathematics strategy use. In M. Carr (Ed.), *Motivation in mathematics* (pp. 89-11). Caskill, NJ: Hampton Press.

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary Educational Psychology, 26*, 330-347.

Carr, M., & Jessup, D.L. (1997). Gender differences in first-grade mathematics strategy use: social and metacognitive influences. *Journal of Educational Psychology, 89*, 318-328.

Carr, M., Jessup, D.L., & Fuller, D. (1999). Gender differences in first-grade mathematics strategy use: Parent and teacher contributions. *Journal for Research in Mathematics Education, 30*, 20-46.

Carretti, B., Cornoldi, C., De Beni, R., & Palladino, P. (2004). What happens to information to be suppressed in working-memory tasks ? Short and long term effects. *Quarterly Journal of Experimental Psychology, 57A*, 1059-1084.

Carrier, L.M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1339-1348.

Case, R. (1985). *Intellectual development: Birth to adulthood*. San Diego: Academic Press.

Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term-memory span. *Journal of Experimental Child Psychology, 33*, 386-404.

Casey, M.B., Nuttall, R.L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33*, 669-680.

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high-school. *Sociology of Education, 67*, 199-215.

Chen, C., & Uttal, D.H. (1988). Cultural values, parents' beliefs, and children's achievement in the United States and China. *Human Development, 31*, 351-358.

Chipman, S.F., Krantz, D.H., & Silver, R. (1992). Mathematics anxiety and science careers among able college women. *Psychological Science, 3*, 292-295.

Chochon, F., Cohen, L., van de Moortele, P.F., & Dehaene, S. (1999). Differential contributions of the left and the right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience, 11*, 617-630.

Cohen, L., & Dehaene, S. (2000). Calculating without reading: Unsuspected residual abilities in pure alexia. *Cognitive Neuropsychology, 17*, 563-583.

Cohen, L., Dehaene, S., Chochon, F., Lehericy, S., & Naccache, L. (2000). Language and calculation within the parietal lobe: A combined cognitive, anatomical and fMRI study. *Neuropsychologia, 38*, 1426-1440.

Compton, B.J., & Logan, G.D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition, 19*, 151-158.

Conway, A.R.A., Cowan, N., Bunting, M.F., Therriault, D.J., & Minkoff, S.R.B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163-183.

Conway, A.R.A., & Engle, R.W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 354-373.

Cooney, J.B., Swanson, H.L., & Ladd, S.F. (1988). Acquisition of mental multiplication skill; Evidence for the transition between counting and retrieval strategies. *Cognition and Instruction, 5*, 323-345.

Cornet, J.A., Seron, X., Deloche, G., & Lories, G. (1988). Cognitive models of simple mental arithmetic – A critical review. *Cahiers de Psychologie Cognitive, 8*, 551-571.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*, 163-191.

Cowan, N. (1995). *Attention and Memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge, UK: Cambridge University Press.

Cowan, N. (1999). An embedded-process model of working memory. In a. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

Craik, F.I.M., Govoni, R., Naveh-Benjamin, M., & Anderson, N.D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*, 159-180.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Dark, V.J., & Benbow, C.P. (1991). Differential enhancement of working memory with mathematical versus verbal precocity. *Journal of Educational Psychology, 83*, 48-60.

Dar-Nimrod, I., & Heine, J. (2006). Exposure to scientific theories affects women's math performance. *Science, 314*, 435.

Davis, H., & Carr, M. (2002). Gender differences in mathematics strategy use. The influence of temperament. *Learning and Individual Differences, 13*, 83-95.

De Brauwer, J., & Fias, W. (2007). *Automatic retrieval in solving simple division problems*. Manuscript under revision.

De Brauwer, J., Verguts, T., & Fias, W. (2006). The representation of multiplication facts: Developmental changes in the problem size, five and tie effects. *Journal of Experimental Child Psychology, 94*, 43-56.

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. *Journal of Educational Psychology, 82*, 359-365.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Dehaene, S. (1997). *The Number Sense*. New York: Oxford University Press.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83-120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33*, 219-250.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three paretial circuits for number processing. *Cognitive neuropsychology, 20*, 487-506.

De Jonge, P., & De Jong, P. (1996). Working memory, intelligence and reading ability in children. *Personality and Individual Differences, 21*, 1007-1020.

Delaney, P.F., Reder, L.M., Staszewski, J.J., & Ritter, F.E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science, 9*, 1-7.

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic – An fMRI study. *Cognitive Brain Research, 18*, 76-88.

Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C.M., Kaufmann, L., Benke, T., & Felber, S. (2005). Learning by strategies and learning by drill – Evidence from an fMRI study. *Neuroimage, 25*, 838-849.

Delgado, A.R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence, 32*, 25-32.

De Rammelaere, S., Stuyven, E., Vandierendonck, A. (1999). The contribution of working memory resources in the verification of simple arithmetic sums. *Psychological Research, 62*, 72-77.

De Rammelaere, S., Stuyven, E., Vandierendonck, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved? *Memory & Cognition, 29*, 267-273.

De Rammelaere, S., & Vandierendonck, A. (2001). Are executive processes used to solve simple mental arithmetic production tasks? *Current Psychology Letters: Behaviour, Brain & Cognition, 5*, 79-89.

Deschuyteneer, M., & Vandierendonck, A. (2005a). Are “input monitoring” and “response selection” involved in solving simple mental arithmetical sums? *European Journal of Cognitive Psychology, 17*, 347-370.

Deschuyteneer, M., & Vandierendonck, A. (2005b). The role of “response selection” and “input monitoring” in solving simple arithmetical products. *Memory & Cognition, 33*, 1472-1483.

Deschuyteneer, M., Vandierendonck, A., & Coeman, P. (2007). *Exploring the role of different executive processes in solving simple mental arithmetical subtractions and divisions*. Manuscript under revision.

Deschuyteneer, M., Vandierendonck, A., & Muyllaert, I. (2006). Does solution of mental arithmetic problems such as 2 + 6 and 3 x 8 rely on the process of “memory updating”? *Experimental Psychology, 53*, 198-208.

De Smedt, B., Ghesquière, P., & Verschaffel, L. (2004). *Working memory and arithmetic abilities in seven and eleven year old children.* Poster presented at the Second European Working Memory Symposium, Beaune, France.

De Smedt, B., Swillen, A., Devriendt, K., Fryns, J.P., Verschaffel, L., & Ghesquière, P. (2006). Mathematical disabilities in young primary school children with Velo-Cardio-Facial Syndrome. *Genetic Counseling, 17, 259-280.*

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386.

De Vos, T. (1992). *Tempo Test Rekenen [Arithmetic Tempo Test]*. Lisse, Switserland: Swets and Zeitlinger.

Dierckx, V., & Vandierendonck, A. (2005). Adaptive strategy application in linear reasoning. In Roberts, M.J., & Newton, E.J. (Eds.). *Methods of thought: Individual differences in reasoning strategies* (pp. 107-182). Hove, UK: Psychology Press.

Donlan, C., Cowan, R., Newton, E.J., & Lloyd, D. (in press). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition*.

Dowker, A. (1996). How important is spatial ability to mathematics? *Behavioral and Brain Sciences, 19*, 251.

Dowker, A. (1997). Young children’s addition estimates. *Mathematical Cognition, 3*, 141-154.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91*, 113-136.

Duverne, S., & Lemaire, P. (2004). Age-related differences in arithmetic problem-verification strategies. *Journal of Gerontology, 59B*, 135-142.

Duverne, S., & Lemaire, P. (2005). Arithmetic split effects reflect strategy selection: An adult age comparative study in addition comparison and verification tasks. *Canadian Journal of Experimental Psychology, 59*, 262-278.

Duverne, S., Lemaire, P., & Michel, B.F. (2003). Alzheimer's disease disrupts arithmetical fact retrieval processes but not arithmetic strategy selection. *Brain and Cognition, 52*, 302-318.

Duverne, S., Lemaire, P., & Vandierendonck, A. (in press). Do working-memory executive components mediate the effects of age on strategy selection or on strategy execution? Insights from arithmetic problem solving. *Psychological Research.*

Eccles, J., Wigfield, A., Harold, R.D., & Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. *Child Development, 64*, 830-847.

Ekstrom, R.B., French, J.W., & Harman, H.H. (1976). *Manual for kit of factor-referenced cognitive tests: 1976*. Princeton, NJ: Educational Testing Service.

El Yagoubi, R., Lemaire, P., & Besson, M. (2003). Different brain mechanisms mediate two strategies in arithmetic: Evidence from Event-Related brain Potentials. *Neuropsychologia, 41*, 855-862.

Engle, R.W. (2001). What is working memory capacity? In H.L. Roediger III, J.S. Nairne, I. Neath, & M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297-314). Washington, DC: American Psychology Association.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Engle, R.W., Cantor, J., & Carullo, J.J. (1992). Individual differences in working memory and comprehension: A test of 4 hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 972-992.

Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge: Cambridge University Press.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R. (1999b). Working memory, STM, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Entwisle, D.R., & Baker, D.P. (1983). Gender and young children's expectations for performance in arithmetic. *Developmental Psychology, 19*, 200-209.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working-memory. *Psychological Review, 102*, 211-245.

Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Ericsson, K.A., & Simon, H.A. (1993). *Protocal analysis: Verbal reports as data. Revised edition*. Cambridge, MA: MIT Press.

Espy, K.A., McDiarmid, M.M., Cwik, M.F., Stalets, M.M., Hamby, A., & Senn, T.F. (2004). The contribution of executive functions to emergent mathematics skills in preschool children. *Developmental Neuropsychology, 26*, 465-486.

Eysenck, M.W., & Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion, 6*, 409-434.

Faust, M.W., Ashcraft, M.H., & Fleck, D.E. (1996). Mathematics anxiety effects in simple and complex addition. *Mathematical Cognition, 2*, 25-62.

Felson, R.B., & Trudeau, L. (1991). Gender differences in mathematics performance. *Social Psychology Quarterly, 54*, 113-126.

Fendrich, D.W., Healy, A.F., & Bourne, L.E. (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Applied Cognitive Psychology: Applications of Cognitive Theories and Concepts* (pp. 111-133). Hillsdale, NJ: Erlbaum.

Fennema, E., Carpenter, T.P., Jacobs, V.R., Franke, M.L., & Levi, L.W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27*, 6-11.

Fleck, D.E. (1994). *Speed-accuracy tradeoffs in mental addition performance: A cognitive treatise on mathematics anxiety*. Unpublished M.A. thesis, Cleveland State University.

French, J. W., Ekstrom, R. B. & Price, I. A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.

Friedman, L. (1995). The space factor in mathematics – Gender differences. *Review of Educational Resources, 65*, 22-50.

Fry, A.F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*, 237-241.

Fry, A.F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*, 1-34.

Fuchs, L.S., Compton, D.L., Fuchs, D., Paulsen, K., Bryant, J.D., & Hamplett, C.L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97, 493-513.*

Fuchs, L.S., Fuchs, D., Compton, D.L., Powell, S.R., Seethaler, M., Capizzi, A.M., Schatschneider, C., & Fletcher, J.M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98, 29-43.*

Fürst, A. J., Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*, 774-782.

Fuson, K.C. (1982). An analysis of the counting-on procedure in addition. In T.H. Carpenter, J.M. Moser, & T.H. Romberg (Eds), *Addition and subtraction: A cognitive perspective* (pp. 67-78). Hillsdale NJ: Lawrence Erlbaum.

Fuson, K.C. (1988). *Children's counting and concepts of number*. New York: Springer.

Galfano, G., Rusconi, E., & Umiltà, C. (2003). Automatic activation of multiplication facts : Evidence from the nodes adjacent to the product. *Quarterly Journal of Experimental Psychology, 56*, 31-61.

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265-281.

Gathercole, S.E., Pickering, S.J., Ambridge, B., Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177-190.

Gavens, N., & Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language, 51*, 644-657.

Geary, D.C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology, 49*, 363-383.

Geary, D.C. (1993). Mathematical disabilities: Cognitive, neuropsychological and genetic components. *Psychological Bulletin, 114*, 345-362.

Geary, D.C. (1994). *Children's mathematical development: Research and practical applications.* Washington, DC: American Psychological Association.

Geary, D.C. (1996). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition, 1996*, 63-93.

Geary, D.C. (1998). *Male, female: The evolution of human sex differences*. Washington, DC: Washington, DC: American Psychology Association.

Geary, D.C. (1999). Sex differences in mathematical abilities: Commentary on the math-fact retrieval hypothesis. *Contemporary Educational Psychology, 24*, 367-274.

Geary, D.C. (2003). Learning disabilities in arithmetic: Problem solving differences and cognitive deficits. In H.L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 199-212). New York: Guilford.

Geary, D.C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4-15.

Geary, D.C., Bow-Thomas, C.C., Fan, L., & Siegler, R.S. (1993a). Even before formal instruction, Chinese children outperform American children in mental addition. *Cognitive Development, 8*, 517-529.

Geary, D.C., Bow-Thomas, C.C., Liu, F., & Siegler, R. (1996a). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67*, 2022-2044.

Geary, D.C., & Brown, S.C. (1991). Cognitive addition – Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled-children. *Developmental Psychology, 27*, 398-406.

Geary, D.C., Brown, S.C., & Samaranayake, V.A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27*, 787-797.

Geary, D.C., & Burlingham-Dubree, M. (1989). External validation of the strategy choice model for addition. *Journal of Experimental Child Psychology, 47*, 175-192.

Geary, D.C., Fan, L., & Bow-Thomas, C.C. (1992). Numerical cognition: Loci of ability differences comparing children from China ant the United States. *Psychological Science, 3*, 180-185.

Geary, D.C., Frensch, P.A., & Wiley, J.G. (1993b). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Journal of Experimental Child Psychology, 47*, 175-192.

Geary, D.C., Hamson, C.O., Chen, G.-P., Fan, L., Hoard, M.K., & Salthouse, T.A. (1997). Computational and reasoning in arithmetic: Cross-generational change in China and the United States. *Psychonomic Bulletin & Review, 4*, 254-263.

Geary, D.C., Hamson, C.O., & Hoard, M.D. (2000a). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77*, 236-263.

Geary, D.C., Hoard, M.K., Byrd-Craven, J., & DeSoto, M.C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*, 121-151.

Geary, D.C., Hoard, M.K., & Hamson, C.O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74*, 213-239.

Geary, D.C., Salthouse, T.A., Chen, G.-P., & Fan, L. (1996b). Are East Asian versus American differences in arithmetic ability a recent phenomenon? *Developmental Psychology, 32*, 254-262.

Geary, D.C., Saults, S.J., Liu, F., & Hoard, M.K. (2000b). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology, 77*, 337-353.

Geary, D.C., & Widaman, K.F. (1987). Individual differences in cognitive arithmetic. *Journal of Experimental Psychology: General, 116*, 154-171.

Geary, D.C., & Widaman, K.F. (1992). Numerical cognition – On the convergence of componential and psychometric models. *Intelligence, 16*, 47-80.

Geary, D.C., Widaman, K.F., & Little, I.D. (1986). Cognitive addition and multiplication: Evidence for a single memory network. *Memory & Cognition, 14*, 478-487.

Geary, D.C., & Wiley, J.G. (1991). Cognitive addition: Strategy choice and speed of processing differences in young and elderly adults. *Psychology & Aging, 6*, 474-483.

Gelman, R., & Gallistel, R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gilles, P.Y., Masse, C., & Lemaire, P. (2001). Individual differences in arithmetic strategy use. *L'Année Psychologique, 101*, 9-32.

Gilmore, C.K. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development, 21*, 301-316.

Goldman, S.R., Mertz, D.L., & Pellegrino, J.W. (1989). Individual differences in extended practice functions and solution strategies for basic addition facts. *Journal of Educational Psychology, 81*, 481-496.

Goldman, S.R., Pellegrino, J.W., & Mertz, D.L. (1988). Extended practice of basic addition facts: Strategy changes in learning-disabled students. *Cognition and Instruction, 5*, 223-265.

Goldstein, J.H. (1994). *Sex differences in toy play and use of video games*. Cambridge, UK: Cambridge University Press.

Green, H.J., Lemaire, P., & Dufau, S. (in press). Eye movement correlates of younger and older adults' strategies for complex addition. *Acta Psychologica*.

Groen, G.J., & Parkman, J.M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329-343.

Grube, D. (2005). *Strategy development in elementary school children's simple addition.* Paper presented at the 11th Meeting of the European Association for Research on Learning and Instruction, Nicosia, Cyprus.

Gruber, O., Indefrey, P., Steinmetz, H., Kleinschmidt, A. (2001). Dissociating neural correlates of cognitive components in mental calculation. *Cerebral Cortex, 11*, 350-359.

Hale, S. (1190). A global developmental trend in cognitive processing speed. *Child Development, 61*, 653-663.

Hamann, M.S., & Ashcraft, M.H. (1985). Simple and complex mental addition across development. *Journal of Experimental Child Psychology, 40*, 49-72.

Hatano, G. (2004). *Cognitive development through participation in practices*. Key note on the 18th Meeting of the International Society for the Study of Behavioral Development, Ghent, Belgium.

Heathcote, D. (1994). The role of visual-spatial working memory in the mental addition of multi-digit addends. *Cahiers de Psychologie Cognitive, 13*, 207-245.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavior Research Methods, Instruments, & Computers, 28*, 427-445.

Hecht, S.A. (1999). Individual solution processes while solving addition and multiplication facts in adults. *Memory & Cognition, 27*, 1097-1107.

Hecht, S.A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory & Cognition, 30*, 447-455.

Hecht, S.A. (in press). Group differences in adult simple arithmetic: Good retrievers, not-so-good retrievers, and perfectionists. *Memory & Cognition*.

Hecht, S.A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology, 86*, 277-302.

Hecht, S.A., Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from 2nd to 5th grades. *Journal of Experimental Child Psychology, 79*, 192-227.

Hegarty, M., Mayer, R.E., & Green, C.E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*, 76-84.

Hegarty, M., Mayer, R.E., & Monk, C.A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*, 18-32.

Hegarty, M., Shah, P., & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition, 28*, 376-385.

Heitz, R.P., Unsworth, N., & Engle, R.W. (2005). Working memory capacity, attention control, and fluid intelligence. In O. Wilhelm & R.W. Engle (Eds.), *Handbook of understanding and measuring intelligence*. London: Sage Publications.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics education, 21*, 33-46.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14*, 251-283.

Hitch, G.J. (1978). Role of short-term memory in mental arithmetic. *Cognitive Psychology, 10*, 302-323.

Hitch, G.J., Cundick, J., Haughey, M., Pugh, R., & Wright, H. (1987). Aspects of counting in children's arithmetic. In J.A. Sloboda and D. Rogers (Eds.), *Cognitive Processes in Mathematics* (pp. 26-41). Oxford: Clarendon Press.

Hitch, G.J., & McAuley, E. (1991). Working memory in children with specific arithmetical difficulties. *British Journal of Psychology, 82*, 375-386.

Hitch, G.J., & Towse, J.N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130*, 184-198.

Hockley, W.E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 598-615.

Hohle, R.H. (1965). Inferred components reaction time as a function of foreperiod duration. *Journal of Experimental Psychology, 69*, 382-386.

Hopkins, S.L., & Lawson, M.J. (2002). Explaining the acquisition of a complex skill: Methodological and theoretical considerations uncovered in the study of simple addition and the moving-on process. *Educational Psychology Review, 14*, 121-154.

Hopkins, S.L., & Lawson, M.J. (2006). The effect counting speed has on developing a reliance on retrieval in basic addition. *Contemporary Educational Psychology, 31*, 208-227.

Hopko, D.R., McNeil, D.W., Lejuez, C.W., Ashcraft, M.H., Eifert, G.H., & Riel, J. (2003). The effects of anxious responding on mental arithmetic and lexical decision task performance. *Anxiety Disorders, 17*, 647-665.

Hoyer, W.J., Cerella, J., & Onyper, S.V. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition, 31*, 1260-1270.

Hutton, U.M.Z., & Towse, J.N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory, 9*, 383-394.

Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics performance: A metaanalysis. *Psychological Bulletin, 107*, 139-155.

Hyde, J.S., Fennema, E., Ryan, M., Frost, L.A., Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A metaanalysis. *Pscyhology of Women Quarterly, 14*, 299-324.

Imbo, I., De Rammelaere, S., & Vandierendonck, A. (2005). New insights in the role of working memory in carry and borrow operations. *Psychologica Belgica, 45*, 101-121.

Imbo, I., Duverne, S., & Lemaire, P. (in press a). Working memory, strategy execution, and strategy selection in mental arithmetic. *Quarterly Journal of Experimental Psychology.*

Imbo, I., & Vandierendonck, A. (2007a). *Practice effects on strategy selection and strategy efficiency in simple mental arithmetic.* Manuscript under revision.

Imbo, I., & Vandierendonck, A. (2007b). *Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: Differences between children and adults?* Manuscript submitted for publication.

Imbo, I., & Vandierendonck, A. (in press b). The role of phonological and executive working-memory resources in simple-arithmetic strategies. *European Journal of Cognitive Psychology.*

Imbo, I., & Vandierendonck, A. (in press c). Do multiplication and division strategies rely on executive and phonological working-memory resources? *Memory & Cognition.*

Imbo, I., & Vandierendonck, A. (in press d). The development of strategy use in elementary-school children: Working memory and individual differences. *Journal of Experimental Child Psychology.*

Imbo, I., Vandierendonck, A., & De Rammelaere, S. (in press e). The role of working memory in the carry operation of mental arithmetic: Number and value of the carry. *Quarterly Journal of Experimental Psychology.*

Imbo, I., Vandierendonck, A., & Rosseel, Y. (in press f). The influence of problem features and individual differences on strategic performance in simple arithmetic. *Memory & Cognition.*

Imbo, I., Vandierendonck, A., & Vergauwe, E. (in press g). The role of working memory in carrying and borrowing. *Psychological Research.*

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstatter, F., Benke, T., Felber, S., Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *Neuroimage, 30*, 1365-1375.

Jackson, N., & Coney, J. (2005). Simple arithmetic processing: The question of automaticity. *Acta Psychologica, 119*, 41-66.

Jacobs, J.E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Education Psychology, 83*, 518-527.

Jacobs, J.E., & Eccles, J.S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63*, 932-944.

Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509-527.

Johnson, E.S. (1984). Sex-differences in problem-solving. *Journal of Educational Psychology, 76*, 1359-1371.

Jordan, N.C., & Montani, T.O. (1997). Cognitive arithmetic and problem solving : A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities, 30*, 624-634.

Jurden, F.H. (1995). Individual differences in working-memory and complex cognition. *Journal of Educational Psychology, 87*, 93-102.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kail, R. (1988). Developmental functions for speeds of cognitive processes. *Journal of Experimental Child Psychology, 45*, 339-364.

Kail, R. (1990). *The development of memory in children*. San Francisco: Freeman.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109*, 490-501.

Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28*, 899-904.

Kail, R., & Hall, L.K. (1999). Sources of developmental change in children's word-problem performance. *Journal of Educational Psychology, 91*, 660-668.

Kail, R., & Hall, L.K. (2001). Distinguishing short-term memory from working memory. *Memory & Cognition, 29*, 1-9.

Kail, R., & Park, Y.S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57*, 281-291.

Kalaman, D.A., & LeFevre, J.-A. (in press). Working memory demands of exact and approximate addition. *European Journal of Cognitive Psychology*.

Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*, 169-183.

Kaufmann, L. (2002). More evidence for the role of the central executive in retrieving arithmetic facts: A case study of severe developmental dyscalculia. *Journal of Clinical and Experimental Neuropsychology, 24*, 302-310.

Kaufmann, L., Lochy, A., Drexler, A., & Semenza, C. (2003). Deficient arithmetic fact retrieval – Storage or access problem? A case study. *Neuropsychologia, 42*, 482-496.

Kawashima, R., Taira, M., Okita, K., Inoue, K., Tajima, N., Yoshida, H., Sasaki, T., Sugiura, M., Watanabe, J., & Fukuda, H. (2004). A functional MRI study of simple arithmetic – A comparison between children and adults. *Cognitive Brain Research, 18*, 225-231.

Kaye, D.B. (1986). The development of mathematical cognition. *Cognitive Development, 1*, 157-170.

Kaye, D.B., de Winstanley, P., Chen, Q., & Bonnefil, V. (1989). Development of efficient arithmetic computation. *Journal of Educational Psychology, 81*, 467-480.

Kellogg, J.S., Hopko, D.R., & Ashcraft, M.H. (1999). The effects of time pressure on arithmetic performance. *Journal of Anxiety disorders, 13*, 591-600.

Kerkman, D.D., & Siegler, R.S. (1997). Measuring individual differences in children's addition strategy choices. *Learning and Individual Differences, 9*, 1-18.

Kimura, D. (1992). Sex differences in the brain. *Scientific American, 267*, 119-125.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Kirk, E.P., & Ashcraft, M.H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 157-175.

Klapp, S.T., Boches, C.A., Trabert, M.L., Logan, G.D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 196-209.

Koshmider, J.W., & Ashcraft, M.H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology, 51*, 53-89.

Kroesbergen, E.H., & Van Luit, J.E.H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education, 24*, 97-114.

Kroesbergen, E., & Van Luit, J. (2005). *Individual differences in young children's mathematics.* Paper presented at the 11th Meeting of the European Association for Research on Learning and Instruction, Nicosia, Cyprus.

Krueger, L.E. (1986). Why 2 x 2 = 5 looks so wrong: On the odd-even rule in product verification. *Memory & Cognition, 14*, 141-149.

Krueger, L.E., & Hallford, E.W. (1984). Why 2 + 2 = 5 looks so wrong: On the odd-even rule in sum verification. *Memory & Cognition, 12*, 171-180.

Kuhn, D., Schauble, L., Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. *Cognition and Instruction, 9*, 285-327.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity. *Intelligence, 14*, 389-433.

Lebiere, C., & Anderson, J.R. (1998). Cognitive arithmetic. In J.R. Anderson & C. Lebiere (Eds.), *The atomic components of thought* (pp. 297-342). Mahwah, NJ: Erlbaum.

Lee, K.-M., & Kang, S.-Y. (2002). Arithmetic operation and working memory: differential suppression in dual tasks. *Cognition, 83*, B63-B68.

LeFevre, J.-A., & Bisanz, J. (1986). A cognitive analysis of number-series problems: Sources of individual differences in performance. *Memory & Cognition, 14*, 287-298.

LeFevre, J.-A., & Bisanz, J. (1987).

LeFevre, J.-A., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic : Evidence for obligatory activation of arithmetic facts. *Memory & Cognition, 16*, 45-53.

LeFevre, J.-A., Bisanz, J., Daley, K.E., Buffone, L., Greenham, S.L., & Sadesky, G.S. (1996a). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General, 125*, 284-306.

LeFevre, J.-A., & Kulak, A.G. (1994). Individual differences in the obligatory activation of addition facts. *Memory & Cognition, 22*, 188-200.

LeFevre, J.-A., Kulak, A.G., & Bisanz, J. (1991). Individual differences and developmental change in the associative relations among numbers. *Journal of Experimental Child Psychology, 52*, 256-274.

LeFevre, J.-A., Kulak, A.G., & Heymans, S.L. (1992). Factors influencing the selection of university majors varying in mathematical content. *Canadian Journal of Behavioural Science – Revue Canadienne des Sciences du Comportement, 24*, 276-289.

LeFevre, J.-A., Lei, Q., Smith-Chant, B., & Mullins, D.B. (2001). Multiplication by eye and by ear for Chinese-speaking and English-speaking adults. *Canadian Journal of Experimental Psychology, 55*, 277-284.

LeFevre, J.-A., & Liu, J. (1997). The role of experience in numerical skill: Multiplication performance in adults from Canada and China. *Mathematical Cognition, 3*, 31-62.

LeFevre, J.-A., & Morris, J. (1999). More on the relation between division and multiplication in simple arithmetic: Evidence for mediation of division solutions via multiplication. *Memory & Cognition, 27*, 803-812.

LeFevre, J.-A., Sadesky, G.S., & Bisanz, J. (1996b). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 216-230.

LeFevre, J.-A., Shanahan, T., & DeStefano, D. (2004). The tie effect in simple arithmetic: An access-based account. *Memory & Cognition, 32,* 1019-1031.

LeFevre, J.-A., Smith-Chant, B.L., Fast, L., Skwarchuk, S.-L., Sargla, E., Arnup, J.S., Penner-Wilger, M., Bisanz, J., & Kamawar, D. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through grade 2. *Journal of Experimental Child Psychology, 93*, 285-303.

Lemaire, P., Abdi, H., & Fayol, M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology, 8*, 73-103.

Lemaire, P., Arnaud, L., Lecacheur, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology and Aging, 19*, 467-481.

Lemaire, P., Barrett, S.E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. *Journal of Experimental Child Psychology, 57*, 224-258.

Lemaire, P., & Fabre, L. (2005). Strategic aspects of human cognition: Implications for understanding human reasoning. In E.J. Newton & M.J. Roberts (Eds.), *Methods of though: Individual differences in reasoning strategies* (pp. 11 – 56). Hove, UK: Psychology Press.

Lemaire, P., & Fayol, M. (1995). When plausibility judgments supersede fact retrieval: The example of the odd-even effect on product verification. *Memory & Cognition, 23*, 34-48.

Lemaire, P., & Lecacheur, M. (2002). Children's strategies in computational estimation. *Journal of Experimental Child Psychology, 82*, 281-304.

Lemaire, P., Lecacheur, M., & Farioli, F. (2000). Children's strategy use in computational estimation. *Canadian Journal of Experimental Psychology, 54*, 141-148.

Lemaire, P., & Machard, I. (2003). *Working memory and adaptiveness of strategy use in arithmetic: Insight from computational estimation in adults and children.* Paper presented at the 13th Conference of the European Society of Cognitive Psychology, Granada, Spain.

Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition, 27*, 364-382.

Lemaire, P., & Siegler, R.S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83-97.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin and Review, 12*, 165-170.

Lépine, R., Bernardin, S., & Barrouillet, P. (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology, 17*, 329-345.

Lépine, R., Roussel, J.L., & Fayol, M. (2003). Procedural resolution or direct retrieval in memory of simple addition and multiplication in children? *Année Psychologique, 103*, 51-80.

Lepola, J., Niemi, P., Kuikka, M., Hannula, M.M. (in press). Cognitive-linguistic skills and motivation as longitudinal predictors of reading and arithmetic achievement: A follow-up study from kindergarten to grade 2. *International Journal of Educational Research.*

Levine, D.R. (1982). Strategy use and estimation ability of college students. *Journal for Research in Mathematics Education, 13*, 350-359.

Lewis, A.B. (1989). Training students to represent arithmetic word problems. *Journal of educational psychology, 81*, 521-531.

Logan, G.D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology, 39*, 367-386.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

Logan, G.D., & Klapp, S.T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 179-195.

Logie, G.D. (1995). *Visuo-spatial working memory*. Hove: Lawrence Erlbaum Associates, Ltd.

Logie, R.H., & Baddeley, A.D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*, 310-326.

Logie, R. H., Gilhooly, K. J., Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*, 395-410.

Lovett, M.C., Reder, L.M., & Lebiere, C. (1999). Modeling working memory in an unified architecture: An ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of maintenance and executive control* (pp. 135-182). Cambridge, UK: Cambridge University Press.

Luchins, A.S. (1942). *Mechanization in problem solving*. Psychological Monographs, 54, No. 248.

Lummis, M., & Stevenson, H.W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26*, 254-263.

Luwel, K., Verschaffel, L., Onghena, P., De Corte, E. (2003). Analysing the adaptiveness of strategy choices using the choice/no-choice method: The case of numerosity judgement. *European Journal of Cognitive Psychology, 15*, 511-537.

Luwel, K., Lemaire, P., & Verschaffel, L. (2005). Children's strategies in numerosity judgment. *Cognitive Development, 20*, 448-471.

Luwel, K., Bulté, I., Onghena, P., & Verschaffel, L. (2007). *Does switching between cognitive strategies entail a cost?* Manuscript in preparation.

Mabbott, D.J., & Bisanz, J. (2003). Developmental change and individual differences in children's multiplication. *Child Development, 74*, 1091-1107.

Martein, R., Kemps, E., & Vandierendonck, A. (1999). The role of working memory in a double span task. *Psychologica Belgica, 39*, 15-27.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*, 107-157.

McLean, J.F., & Hitch, G.J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240-260.

Meece, J.L., Wigfield, A., Eccles, J.S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*, 60-70.

Menon, V., Rivera, S.M., White, C.D., Eliez, S., Glover, G.H., & Reiss, A.L. (2000). Functional optimization of arithmetic processing in perfect performers. *Cognitive Brain Research, 9*, 343-345.

Metcalfe, A.W.S., & Campbell, J.I.D. (in press). The role of cue familiarity in adults' strategy choices for simple addition. *European Journal of Cognitive Psychology.*

Miller, K., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. *Journal of Experimental Psychology, 10*, 46-60.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General, 130*, 163-168.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Mulhern, G., & Wylie, J. (2004). Changing levels of numeracy and other core mathematical skills among psychology undergraduates between 1992 and 2002. *British Journal of Psychology, 95*, 355-370.

Nelder, J.A., & Mead, R. (1965). A simplex algorithm for function minimization. *Computer Journal, 7*, 308-313.

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Noël, M.-P., Désert, M., Aubrun, A., & Seron, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory & Cognition, 29*, 34-42.

Noël, M.-P., Seron, X., & Trovarelli, F. (2004). Working memory as a predictor of addition skills and addition strategies in children. *Current Psychology of Cognition, 22*, 3-25.

Onyper, S.V., Hoyer, W.J., & Cerella, J. (2006). Determinants of retrieval solutions during cognitive skill training: Source confusions. *Memory & Cognition, 34*, 538-549.

Pashler, H. (1993). Dual-task interference and elementary mental mechanisms. *Attention and Performance, 14*, 245-264.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Passolunghi, M.C., Cornoldi, C., & De Liberto, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition, 27*, 779-790.

Passolunghi, M.C., & Pazzaglia, F. (2005). A comparison of updating processes in children good or poor in arithmetic word problem-solving. *Learning & Individual Differences, 15*, 257-269.

Passolunghi, M.C., & Siegel, L.S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80*, 44-57.

Passolunghi, M.C., & Siegel, L.S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348-367.

Passolunghi, M.C., Vercelloni, B., & Schadee, H. (in press). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development.*

Pauli, P., Bourne, L.E., & Birbaumer, N. (1998). Extensive practice in mental arithmetic and practice transfer over a ten-month retention interval. *Mathematical Cognition, 4*, 21-46.

Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, N., Rickard, T.C., Yaroush, R.A., & Bourne, L.E. (1994). Brain potentials during mental arithmetic – Effects of extensive practice and problem difficulty. *Cognitive Brain Research, 2*, 21-29.

Pauli, P., Lutzenberger, W., Birbaumer, N., Rickard, T.C., Bourne, L.E. (1996). Neurophysiological correlates of mental arithmetic. *Psychophysiology, 33*, 522-529.

Penner-Wilger, M., Leth-Steensen, C., & LeFevre, J.-A. (2002). Decomposing the problem-size effect: A comparison of response time distributions across cultures. *Memory & Cognition, 30*, 1160-1167.

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of Arabic number processing, numerical comparison, and simple addition: A PET study. *Journal of Cognitive Neuroscience, 12*, 461-479.

Pesenti, M., Zago, L., Crivello, F., Mellet, E., Samson, D., Duroux, B., Seron, X., Mazoyer, B., Tzourio-Mazoyer, N. (2001). Mental calculation in a prodigy is sustained by right prefrontal and medial temporal areas. *Nature Neuroscience, 4*, 103-107.

Phillips, W.A., & Christie, D.F.M. (1977). Interference with visualization. *Quarterly Journal of Experimental Psychology, 29*, 637-650.

Pickering, S.J. (2001). The development of visuo-spatial working memory. *Memory, 9*, 423-432.

Qin, Y., Carter, C.S., Silk, E.M., Stenger, V.A., Fissell, K., Goode, A., & Anderson, J.R. (2004). The change of the brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 5686-5691.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*, 137-157.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446-461.

Ratcliff, R., & Murdock, B.B. (1976). Retrieval processes in recognition memory. *Psychological review, 83*, 190-214.

Reder, L.M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90-138.

Reder, L.M., & Ritter, F.E. (1992). What determines initial feeling of knowing – Familiarity with questions terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 435-451.

Richardson, F.C., & Suinn, R.M. (1972). Mathematics anxiety rating scale – Psychometric data. *Journal of Counseling Psychology, 19*, 551-554.

Rickard, T.C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126,* 288-311.

Rickard, T.C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 250-257.

Rickard, T.C., & Bourne, L.E. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1281-1295.

Rickard, T.C., Healy, A.F., & Bourne, L.E. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 1139-1153.

Rittle-Johnson, B., & Alibali, M.W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189.

Rittle-Johnson, B., Siegler, R.S. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development, 70*, 332-348.

Rittle-Johnson, B., Siegler, R.S., & Alibali, M.W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346-362.

Rivera, S.M., Reiss, A.L., Eckert, M.A., Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*, 1779-1790.

Robert, M., & Savoie, N. (2006). Are there gender differences in verbal and visuospatial working-memory resources? *European Journal of Cognitive Psychology, 18*, 378-397.

Robinson, K.M. (2001). The validity of verbal reports in children's subtraction. *Journal of Educational Psychology, 93*, 211-222.

Robinson, K.M., Arbuthnott, K.D., & Gibbons, K.A. (2002). Adults' representations of division facts: A consequence of learning history? *Canadian Journal of Experimental Psychology, 56*, 302-309.

Robinson, K.M., Arbuthnott, K.D., Rose, D., McCarron, M.C., Globa, C.A., & Phonexay, S.D. (2006). Stability and change in children's division strategies. *Journal of Experimental Child Psychology, 93*, 224-238.

Rohrer, D., & Pashler, H. (2003). Concurrent task effect on memory retrieval. *Psychonomic Bulletin and Review, 10, 96-103.*

Rohrer, D., Pashler, H., & Etchegaray, J. (1998). When two memories can and cannot be retrieved concurrently. *Memory & Cognition, 26*, 731-739.

Romero, S.G., Rickard, T.C., & Bourne, L.E. (2006). Verification of multiplication facts: An investigation using retrospective protocols. *American Journal of Psychology, 119*, 87-120.

Roussel, J.L., Fayol, M., & Barrouillet, P. (2002). Procedural vs. direct retrieval strategies in arithmetic : A comparison between additive and multiplicative problem solving. *European Journal of Cognitive Psychology, 14*, 61-104.

Royer, J.M., Tronsky, L.N., Chan, Y., Jackson, S.J., & Marchant, H. III. (1999a). Math-fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemporary Educational Psychology, 24*, 181-266.

Royer, J.M., Tronsky, L., Marchant, H. III., & Stanley, J.J. (1999b). Reply to the commentaries on the math-fact retrieval hypothesis. *Contemporary Educational Psychology, 24*, 286-300.

Rusconi, E., Galfano, G., Speriani, V., & Umiltà, C. (2004). Capacity and contextual constraints on product activation: Evidence from task-irrelevant fact retrieval. *Quarterly Journal of Experimental Psychology, 57*, 1485-1511.

Rusconi, E., Galfano, G., Rebonato, E., & Umiltà, C. (2006). Bidirectional links in the network of multiplication facts. *Psychological Research, 70*, 32-42.

Rusconi, E., Priftis, K., Rusconi, M.L., & Umiltà, C. (2006). Arithmetic priming from neglected numbers. *Cognitive Neuropsychology, 23*, 227-239.

Russo, J.E., Johnson, E.J., & Stephens, D.L. (1989). The validity of verbal protocols. *Memory & Cognition, 17*, 759-769.

Salamé, P., & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21*, 150-164.

Salthouse, T.A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2*, 179-183.

Salthouse, T.A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychologica, 79*, 155-170.

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Schunn, C.D., Reder, L.M., Nhouyvanisvong, A., Richards, D.R., & Stroffolino, P.J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 3-29.

Seitz, K., & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology, 12*, 552-570.

Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beiträge, 44*, 275-302.

Seyler, D.J., Kirk, E.P, & Ashcraft, M.H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1339-1352.

Shanahan, T., Lucidi, A., LeFevre, J.-A., & Cestari, V. (2005). *The whole or the sum of the parts: Cultural biases in addition strategies.* Poster presented at the 15th Annual meeting of the Canadian Society for Brain, Behavior, and Cognitive Science at Montréal, Québec (Canada).

Shrager, J., & Siegler, R.S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405-410.

Siegel, L.S., & Linder, B.A. (1984). Short-term memory processes in children with reading and arithmetic learning disabilities. *Developmental Psychology, 20*, 200-207.

Siegel, L.S., & Ryan, E.B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980.

Siegler, R.S. (1986). Unities across domains in children's strategy choices. In M. Perlmutter (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 19, pp. 1-48). Hillsdale, NJ: Erlbaum

Siegler, R.S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250-264.

Siegler, R.S. (1988a). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*, 833-851.

Siegler, R.S. (1988b). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General, 117*, 258-275.

Siegler, R.S. (1989). Hazards of mental chronometry: An example from children's subtraction. *Journal of Educational Psychology, 81*, 497-506.

Siegler, R.S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28*, 225-273.

Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R.S. (1999). Strategic development. *Trends in Cognitive Sciences, 3*, 430-435.

Siegler, R.S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior, 33*, 1-42.

Siegler, R.S., & Jenkins, E.A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

Siegler, R.S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General, 126*, 71-92.

Siegler, R.S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T.J. Simon & G.S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Siegler, R.S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-294). Hillsdale, NJ: Erlbaum.

Siegler, R.S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127*, 377-397.

Smith-Chant, B.L., & LeFevre, J.-A. (2003). Doing als they are told and telling like it is: Self-reports in mental arithmetic. *Memory & Cognition, 31*, 516-528.

So, Y., Hwang, S.-H., Lee, W., Han, C.-W., Lim, K.-R., Woo, Y.-K., Lee, S.-Y., Back, S.-H., Lee, J.-M., Kang, E., & K., S. (2006). *Neuroanatomical correlates of individual differences in arithmetic ability*. Poster presented at the 12th Meeting of Human Brain Mapping, Florence, Italy.

Sohn, M.-H., Goode, A., Koedinger, K.R., Stenger, V.A., Fissell, K., Carter, C.S., & Anderson, J.R. (2004). Behavioral equivalence, but not neural equivalence – Neural evidence of alternative strategies in mathematical thinking. *Nature Neuroscience, 7*, 1193-1194.

Spearman, C. (1927). *The abilities of man*. London: MacMillan.

Spelke, E., & Dehaene, S. (1999). Biological foundations of numerical thinking: response to T.J. Simon. *Trends in Cognitive Science, 3*, 365-366.

Spelke, E.S., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition, 78*, 45-88.

Stanescu-Cosson, R., Pinel, P., van de Moortele, P.F., Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Understanding dissociations in dyscalculia – A brain imaging study of the impact of number size on the cerebral networks for exact and approximate calculation. *Brain, 123*, 2240-2255.

Stazyk, E.H., Ashcraft, M.H., & Hamann, M.S. (1982). A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 320-335.

Steel, S., & Funnell, E. (2001). Learning multiplication facts: A study of children taught by discovery methods in England. *Journal of Experimental Child Psychology, 79*, 37-55.

Stevenson, H.W., Chen, C., & Lee, S. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science, 259*, 53-58.

Stevenson, H.W., Lee, S., Chen, C., Lummis, M., Stigler, J., Fan, L., & Ge, F. (1990). Mathematics achievement of children in China and the United States. *Child Development, 61*, 1053-1066.

Stevenson, H.W., Lee, S., & Stigler, J.W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 231*, 691-699.

Stigler, J.W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*, 12-17.

Stipek, D.J., & Gralinski, J.H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology, 83*, 361-371.

Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 455-477). New York: Elsevier Science Publishing.

Suppes, P., Cohen, M., Laddaga, R., Anliker, J., & Floyd, R. (1983). A procedural theory of eye movements in doing arithmetic. *Journal of Mathematical Psychology, 27*, 341-369.

Svenson, O. (1985). Memory retrieval of answers of simple additions as reflected in response latencies. *Acta Psychologica, 59*, 285-304.

Svenson, O., Hedenborg, M.L., & Lingman, L. (1976). On children's heuristics for solving additions. *Scandinavian Journal of Educational Research, 20*, 161-173.

Svenson, O., & Sjöberg, K. (1981). Subitizing and counting processes in young children. *Scandinavian Journal of Psychology, 19*, 247-250.

Svenson, O., & Sjöberg, K. (1982). Solving simple subtractions during the first three school years. *Journal of Experimental Education, 50*, 91-100.

Svenson, O., & Sjöberg, K. (1983). Evolution of cognitive processes for solving simple additions during the first three school years. *Scandinavian Journal of Psychology, 24*, 117-124.

Swanson, H.L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology, 56*, 87-114.

Swanson, H.L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory & Cognition, 32*, 648-661.

Swanson, H.L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology, 93*, 239-264.

Swanson, H.L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*, 471-491.

Swanson, H.L., & Kim, K. (in press). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence*.

Swanson, H.L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*, 294-321.

Szmalec, A., Vandierendonck, A., & Kemps, E. (2005). Speeded Decision-Making as a Component of Executive Functioning. *Memory & Cognition, 33*, 531-541.

Temple, C.M., & Sherwood, S. (2002). Representation and retrieval of arithmetical facts: Developmental difficulties. *The Quarterly Journal of Experimental Psychology, 55A*, 733-752.

Thevenot, C., Barrouillet, P., & Fayol, M. (2001). Algorithmic solution of arithmetic problems and operands-answer associations in long-term memory. Quarterly *Journal of Experimental Psychology, 54*, 599-611.

Thevenot, C., Barrouillet, P., & Fayol, M. (2004). Mental representation and procedures in arithmetic word problems: The effect of the position of the question. *Année Psychologique, 104*, 683-699.

Thevenot, C., Fanget, M., & Fayol, M. (2005). *Arithmetic procedures or retrieval? An operand recognition paradigm*. Poster presented at the NUMBRA/ESCoP Summer School, Erice, Sicilia (Italy).

Thevenot, C., Fanget, M., & Fayol, M. (in press). Retrieval or nonretrieval strategies in mental arithmetic? An operand-recognition paradigm. *Memory & Cognition*.

Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *Quarterly Journal of Experimental Psychology, 58A*, 1311-1323.

Thevenot, C., Oakhill, J. (2006). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology, 18*, 756-775.

Thimbleby, H. (2000). Calculators are needlessly bad. *International Journal of Human-Computer Studies, 52*, 1031-1069.

Thibodeau, M.H., LeFevre, J.-A., & Bisanz, J. (1996). The extension of the interference effect to multiplication. *Canadian Journal of Experimental Psychology – Revue Canadienne de Psychologie Experimentale, 50*, 393-396.

Thorndike, E. (1922). *The psychology of arithmetic*. New York: MacMillan.

Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2002). Strategic competence: Applying Siegler's theoretical and methodological framework to the domain of simple addition. *European Journal of Psychology of Education, 17*, 275-291.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2004a). Strategic aspects of simple addition and subtraction: the influence of mathematical ability. *Learning and Instruction, 14*, 177-195.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2004b). Strategy development in children with mathematical disabilities: Insights from the Choice/No-Choice method and the Chronological-Age/Ability-Level-Match Design. *Journal of Learning Disabilities, 37*, 119-131.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2005a). Simple addition strategies in a first-grade class with multiple strategy instruction. *Cognition and Instruction, 23, 1-21.*

Torbeyns, J., Verschaffel, L., Ghesquière, P. (2005b). *Split or jump? Children's strategy choices in the number domain up to 100.* Paper presented at the 11th Meeting of the European Association for Research on Learning and Instruction, Nicosia, Cyprus.

Touron, D.R., & Hertzog, C. (2004). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition, 32*, 298-310.

Touron, D.R., Hoyer, W.J., & Cerella, J. (2004). Cognitive skill learning : Age-related differences in strategy shifts and speed of component operations. *Psychology and Aging, 19*, 565-580.

Trbovich, P.L., & LeFevre, J.-A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition, 31*, 738-745.

Tuholski, S.W., Engle, R.W., & Baylis, G.C. (2001). Individual differences in working memory capacity and enumeration. *Memory & Cognition, 29*, 484-492.

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent. *Journal of Memory and Language, 28*, 127-154.

Van der Goten, K., & Vandierendonck, A. (1997). Software to calculate measures of sensitivity and response bias based on detection theory and threshold theory. *Behavior Research Methods, Instruments, & Computers, 29*, 461-463.

van der Sluis, S., de Jong, P.F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87*, 239-266.

van der Sluis, S., de Jong, P.F., & van der Leij, A. (in press). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*.

van der Sluis, S., Posthuma, D., Dolan, C.V., de Geus, E.J.C., Colom, R., & Boomsma, D.I. (2006). Sex differences on the Dutch WAIS-III. *Intelligence, 34*, 273-289.

Venkatraman, V., Ansari, D., Chee, M.W.L. (2005). Neural correlates of symbolic and non-symbolic arithmetic. *Neuropsychologia, 43*, 744-753.

Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition, 33*, 1-16.

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience and Biobehavioral Reviews, 26*, 849-857.

Vermeer, H.J., Boekaerts, M., & Seegers, G. (2000). Motivational and gender differences: Sixth-grade students' mathematical problem-solving behavior. *Journal of Educational Psychology, 92*, 308-315.

Verschaffel, L., De Corte, E., Gielen, I., & Struyf, E. (1994). Clever rearrangement strategies in children's mental arithmetic: A confrontation of eye-movement data and verbal protocols. In J.E.H.V. Luit (Ed.), *Research on learning and instruction of mathematics in kindergarten and primary school* (pp. 153-180). Doetincham, The Netherlands; Rapallo, Italy: Graviant Publishing Company.

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*, 85-94.

Verschaffel, L., Janssens, S., & Janssen, R. (2005). The development of mathematical competence in Flemish preservice elementary school teachers. *Teaching and Teacher Education, 21*, 49-63.

Wechsler, D. (1986). *Manual for the Wechsler Intelligence Scale for Children – Revised*. Harcourt Test Publishers.

Whalen, J., McCloskey, M., Lindemann, M., & Bouton, G. (2002). Representing arithmetic table facts in memory: Evidence from acquired impairments. *Cognitive Neuropsychology, 19*, 505-522.

Whetstone, T. (1998). The representation of arithmetic facts in memory: Results from retraining a brain-damaged patient. *Brain and Cognition, 36*, 290-309.

Widaman, K.F., Geary, D.C., Cormier, P., & Little, T.D. (1989). A componential model for mental addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 898-919.

Widaman, K.F., & Little, T.D. (1992). The development of skill in mental arithmetic: An individual differences perspective. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 189-253). Amsterdam: Elsevier Science Publishers.

Widaman, K.F., Little, T.D., Geary, D.C., & Cormier P. (1992). Individual differences in the development of skill in mental addition: Internal and external validation of chronometric models. *Learning and Individual Differences, 4*, 167-213.

Wilson, J.T.L., Scott, J.H., & Power, K.G. (1987). Developmental differences in the span of visual memory for pattern. *British Journal of Developmental Psychology, 5*, 249-255.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193.

Zago, L., Pesenti, M., Mellet, E., Crivello, F., Mazoyer, B., Tzourio-Mazoyer, N. (2001). Neural correlates of simple and complex mental calculation. *Neuroimage, 13*, 314-327.

Zbrodoff, N.J. (1995). Why is 9 + 7 harder than 2 + 3? Strength and interference as explanations of the problem-size effect. *Memory & Cognition, 23*, 689-700.

Zbrodoff, N.J. (1999). Effects of counting in alphabet arithmetic: Opportunistic stopping and priming of intermediate steps. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 299-317.

Zbrodoff, N.J., & Logan, G.D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115*, 118-130.

Zbrodoff, N.J., & Logan, G.D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 83-97.